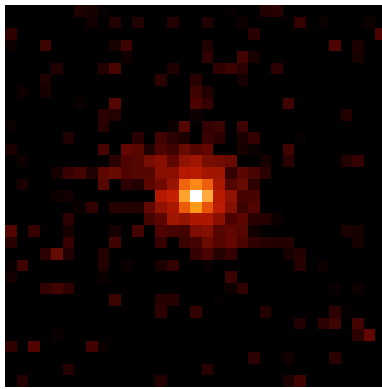


A Mid-Infrared Camera for the Next Generation Space Telescope

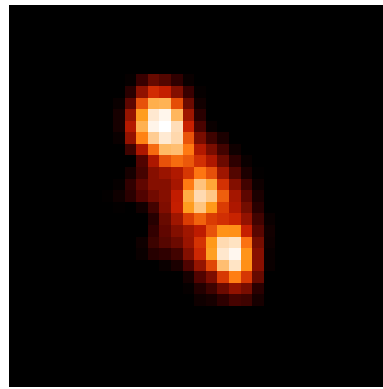
Gene Serabyn, P.I.

Jet Propulsion Laboratory

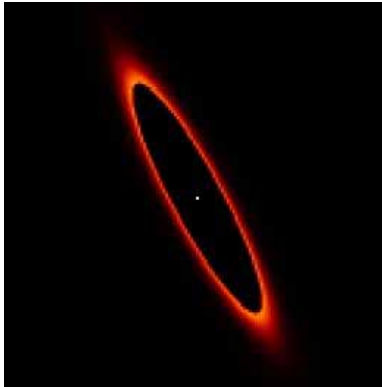
September 1, 1999



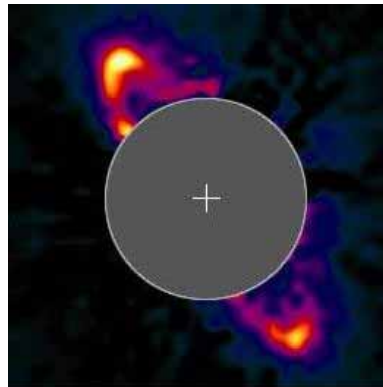
12.5 μm



24.5 μm



Model



HST/NICMOS

HR4796

A Report to the National Aeronautics and Space Administration

Contributing authors:

- L. Armus, SIRTf Science Center, California Institute of Technology
- D. E. Backman, Franklin and Marshall College and NASA Ames Research Center
- M. Barsony, Harvey Mudd College and Jet Propulsion Laboratory
- C. A. Beichman, Jet Propulsion Laboratory
- R. C. Bowman, Jet Propulsion Laboratory
- D. Crumb, Swales Aerospace
- Z. Haiman, Fermilab, and Princeton University
- L. C. Ho, Observatories of the Carnegie Institution of Washington
- D. W. Koerner, University of Pennsylvania
- C. A. Lindensmith, Jet Propulsion Laboratory
- S. A. Macenka, Jet Propulsion Laboratory
- M. A. Malkan, University of California, Los Angeles
- S. P. Oh, Princeton University
- G. S. Orton, Jet Propulsion Laboratory
- M. E. Ressler, Jet Propulsion Laboratory
- N. A. Raouf, Jet Propulsion Laboratory
- E. Serabyn, Jet Propulsion Laboratory and California Institute of Technology
- D. L. Shupe, SIRTf Science Center, California Institute of Technology
- E. F. van Dishoeck, Sterrewacht Leiden
- M. W. Werner, Jet Propulsion Laboratory
- B. Whitney, Space Science Institute
- R. A. Woodruff, Ball Aerospace & Technologies Corp.
- C. Xu, Infrared Analysis and Processing Center, California Institute of Technology

Acknowledgments are also due to our many colleagues who willingly contributed figures and discussions of the issues involved in this report.

The research described herein was carried out in part by the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

Reference herein to any specific commercial product, process, or service by trade name, trademark, or manufacturer, does not constitute or imply its endorsement by the United States Government or the Jet Propulsion Laboratory, California Institute of Technology.

1 Executive Summary

Extending the capabilities of NGST into the mid-infrared (MIR) would vastly enhance the ability of the observatory to resolve key scientific questions. The areas most directly impacted by an MIR capability on NGST include the genesis and evolution of galaxies, the nature of luminous galactic nuclei, the star formation process, the life cycle of dust grains, the evolution of circumstellar particulate disks, and the detection of thermal emission from planets around nearby stars. These topics can be summarized as the formation of galaxies, of stars, of planetary systems, and of the building blocks of life. The range of topics which can be tackled with an MIR capability on NGST is thus quite broad, and it is crucial to note that several of these topics can *only* be resolved with an MIR-capable NGST. This report details the wide range of scientific issues for which an MIR capability on NGST is crucial, presents a flexible design concept for an MIR camera/spectrograph, and explores the critical area of cooling detector arrays to operating temperatures below the NGST passive cooling regime. In order to fully separate camera and cryocooler issues and costs, cryocoolers are addressed in a self-contained appendix.

In this report, we take the MIR as the 5–30 μm wavelength range (sensitivities degrade to longer wavelengths, so those are not considered here). Within this range, several detector possibilities exist on the NGST timescale, with varying cutoff wavelengths and cooling needs, so we also seek to elaborate exactly which MIR wavelengths are most crucial for the various categories of science, to enable the most appropriate and cost-effective capabilities to be implemented. Generally speaking, the astrophysical investigations which an MIR capability enables fall into two categories: rest-frame MIR emission processes in the local universe, and intrinsically shorter-wavelength emission which cosmic expansion has redshifted into the MIR.

Short MIR wavelengths, $\approx 5\text{--}10\ \mu\text{m}$, are especially important to high-redshift (high- z) science for several reasons (note: $z \equiv \Delta\lambda/\lambda$). First, $5\ \mu\text{m}$ is not a true physical limit from the point of view of cosmology: cosmic expansion redshifts a large part of a galaxy's familiar rest-frame near-infrared (NIR) emission into the MIR by $z \sim 3$, and even the $\text{H}\alpha$ line, the strongest emission feature in the visual spectrum of most galaxies, shifts into the MIR by $z = 6.6$. Second, since a significant fraction of high- z galaxies are likely to be dusty, the MIR (i.e. the redshifted NIR) will be vital for peering through the obscuring dust, in order to measure true star-formation rates and the ambient star-forming environments at early epochs. Third, in contrast to the NIR, the majority of the 5–10 μm spec-

tral region (5–8 μm) is completely inaccessible from the ground. Finally, the precursor SIRTf mission will have no high resolution spectroscopic capability for $\lambda < 10\ \mu\text{m}$. The 5–10 μm spectral band is thus both a natural extension to NGST's core cosmological mission concept, and an area which will remain quite underdeveloped spectroscopically even a decade from now (except for inroads made by SOFIA). It is also a natural extension to lower resolution SIRTf imaging at 3.5–8 μm . Thus a sensitive MIR camera/spectrometer covering at least the 5–10 μm range, with a spatial resolution of order $0.12 - 0.25''$ will enable a revolutionary step forward in high- z studies. NGST's large mirror will provide an MIR spatial resolution comparable to that of the Hubble Space Telescope in the NIR, the VLA in the radio, and the ALMA (Atacama Large Millimeter Array) at millimeter wavelengths, allowing a detailed comparison of star-formation and galactic structure tracers over seven decades in frequency.

However, high-redshift science is by no means the whole story, as several areas of “galactic” science are equally, if not more, promising for investigation by NGST in the MIR. The processes regulating star formation, and the origin and evolution of particulate disks and planetary systems around nearby stars are both topics for which longer wavelength MIR observations will be essential, because thermal emission from dust grains cooler than room temperature rises rapidly to longer MIR wavelengths. For example, the spectacular dust ring seen around the star HR4796 (cover figure) was discovered by means of its thermal MIR emission, and only later coronagraphically imaged by the HST in the NIR. Such discoveries are likely to follow in rapid succession, as fainter disks and more obscured star-formation sites are systematically discovered and studied. SIRTf will lay the groundwork in this area, as the far-infrared (50–100 μm) will be a very sensitive tracer of faint dust disks. Although SIRTf will be ideally suited to the discovery of such systems, its limited spatial resolution will be inadequate for detailed studies of the structures and evolutionary states of these disks (while ALMA will likely be sensitive to the cooler, larger-scale emission regions). NGST can excel in these areas, but only if it operates out to $\approx 30\ \mu\text{m}$, its longest feasible MIR wavelength, and only if it includes an MIR spectrographic capability. In particular, as ISO has shown, the constituents of dusty disks can be identified by means of their MIR vibrational bands. The vastly higher spatial resolution of NGST, in comparison with both ISO and SIRTf, can then be used to go the next step: by mapping the distributions of the minerals present in circumstellar disks, the different thermal and chemical regimes present can be minerologically delineated. Even higher spectral

resolution ($R = \lambda/\Delta\lambda \approx 1000 - 3000$) observations of the H_2 molecule can also be used to trace the gradual disappearance of the gas phase as disks evolve into planetary systems. The combination of high spatial and spectral resolution which NGST can provide can thus revolutionize our view of the formation and evolution of solar systems, and NGST may well see in action the processes which ruled the development of our present day solar system.

A brief report such as this cannot address every aspect of MIR science made possible with an MIR-capable NGST. In the following we therefore attempt to provide a coherent framework wherein most of the foreseen science can be captured. To this end we have identified seven topics in which MIR NGST observations are crucial:

- Finding and identifying high- z MIR-bright galaxies, in order to determine the true star-formation rate at early epochs.
- Using MIR spectroscopy to evaluate the galactic environment in early star-formation sites (including a determination of the role of central accretion sources).
- Establishing the nature of the stellar populations present at early epochs.
- Delineating the chemical lifecycle of particulate dust, from its origins in the outflowing embers of dying stars, to its incorporation into new planetary systems.
- Peering deep into the hearts of cold molecular clouds in order to trace the process of protostar formation.
- Examining dust disks around nearby stars, in order to place our solar system into an evolutionary context.
- Surveying the local neighborhood for the coldest brown dwarfs and for extrasolar planets.

Section 2 addresses each of these seven topics in turn. The discussion makes clear that the “discovery space” of NGST can be very significantly enhanced with an MIR capability, even allowing for the flood of information provided by precursor and coeval missions such as SOFIA, SIRTf, ALMA, and large ground-based optical, infrared, and submillimeter telescopes.

2 Scientific Goals

2.1 The Genesis and Evolution of Galaxies: Unveiling the Hidden Universe with an MIR NGST

One of the grand scientific themes of NGST, and of NASA’s “Origins” program as a whole, is understanding the genesis and evolution of galaxies, the main building blocks of the Universe. The task of cosmology in the coming decades will be to observe this cosmic history directly, from $z = 1$ to $z \geq 10$. Operationally, this will require studying the conversion of interstellar matter into stars in extremely distant objects. This is the fundamental cosmic event which leads to the formation of stellar/planetary systems, and to the synthesis of all elements heavier than Helium. To accomplish its mission of understanding the first generation of stars and the birth of galaxies, NGST will need to observe beyond $\lambda = 5 \mu\text{m}$.

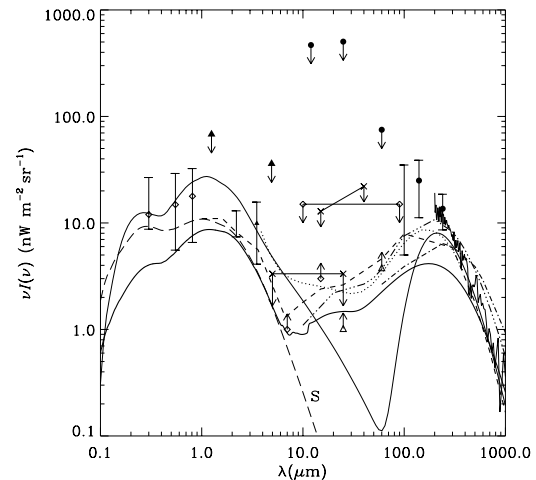


Figure 1: Extragalactic background spectrum from the UV to the mm. The long-dashed line (S) represents a dust-free model with a cosmic star-formation rate as deduced from UV-optical observations alone. (Courtesy E. Dwek)

In the local Universe, regions of high star-formation activity are often dusty, because the youngest, most massive stars are found close to the molecular clouds out of which they formed. A large fraction of the stellar UV/optical luminosity is therefore absorbed by, and heats up, dust grains in the surrounding clouds. As a result, galaxies with the highest star formation rates tend to emit progressively larger fractions of their luminosities in the thermal IR. This reaches extreme conditions in Ultraluminous Infrared Galaxies (ULIRGs) wherein more than 90% of the luminosity emerges in the MIR and far-infrared (FIR).

The discovery of dusty galaxies at high- z with SCUBA and ISO suggests that large populations might remain undiscovered in UV-visual surveys, and that perhaps only a minority of the UV luminosity generated by star formation at high redshifts is detected at all. Indeed, up to half of the expected energy release by nucleosynthesis throughout cosmic time reaches us at FIR and longer wavelengths (Fig. 1), implying that star-formation rates inferred from UV and optical observations alone underestimate the true star-formation rates by at least a factor of two (Dwek *et al.* 1998). Thus, the obscuring effects of dust must be taken into account at high- z at least as much as in the local Universe, and so an NGST limited solely to shorter NIR wavelengths might completely miss much of the star formation in the Universe.

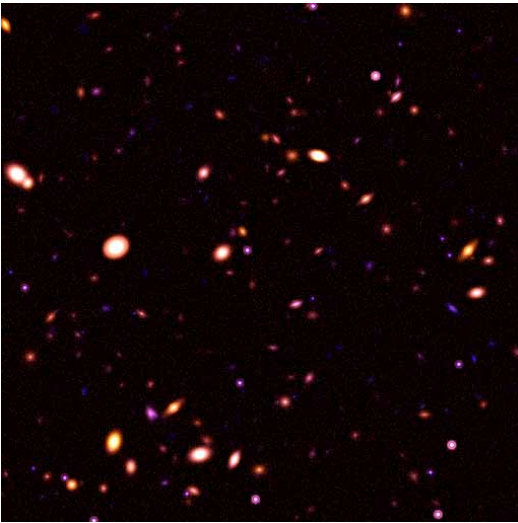


Figure 2: Simulated 10^5 sec MIR integration with a $1k \times 1k$ detector array covering $70''$. Blue represents $10 \mu m$, and red $25 \mu m$. The blue color of most galaxies indicates that they are detected most strongly at $10 \mu m$. Galaxy sizes follow the HDF-South data (Gardner *et al.* 1999). In 10^5 sec, the field contains 850 galaxies at $10 \mu m$ (to $10\text{-}\sigma = 17$ nJy) and 145 galaxies at $25 \mu m$ (to 400 nJy).

To quantify what an MIR-capable NGST would see, we simulated expected 10 and $25 \mu m$ galaxy counts with the empirical model developed for WIRE (Xu *et al.* 1998). The model starts with the properties of galaxies detected by IRAS, and computes fluxes and colors at higher z . The model assumes luminosity evolution increasing as $(1+z)^{3.5}$ to $z=2$, and then no evolution for $z>2$. With these conservative assumptions, a 10^5 sec NGST image will detect about 850 galaxies at $10 \mu m$ and 150 galaxies at $25 \mu m$ in a $70''$ field (Fig. 2). The peak of the red-shift dis-

tribution is at $z \sim 2$ (Fig. 3), and even without evolution at higher z 's, there will be a significant high- z tail: over 300 galaxies at $10 \mu m$ and 40 at $25 \mu m$ with $z \geq 4$. New populations of dusty objects at high- z will thus be easily detectable in even a single deep NGST exposure. Simultaneous observation of both passbands would of course optimize efficiency.

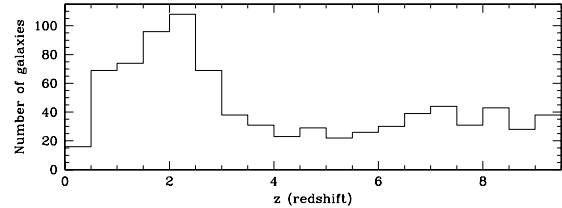


Figure 3: Redshift distribution of the simulated galaxy field of Fig. 2. Note the significant number of objects at $z > 4$.

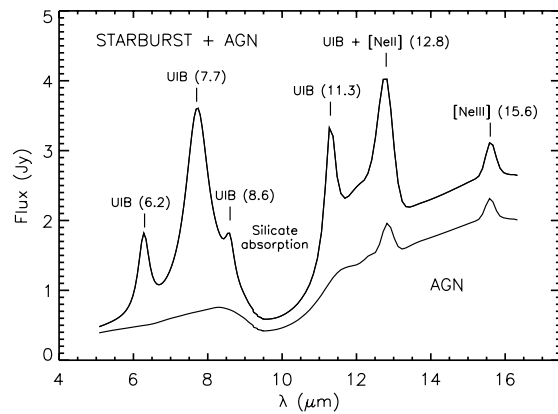
With multi-wavelength observations, the high- z candidates may be identified as Lyman-limit or Gunn-Peterson dropouts, in a manner similar to the detection of Lyman-break objects at $z \sim 3$. Fig. 4 shows the number of high- z objects detectable per field-of-view beyond redshifts of $z=5, 7$, and 10 , in both the NIR and MIR, for the higher star-formation efficiency model of Haiman & Loeb (1998). Although the number of expected galaxies is high in both bands, the MIR is optimal for detecting young quasars (assuming the average quasar spectrum of the Elvis *et al.* 1994 sample, which has an absolute intrinsic flux increasing with λ to the FIR). Of course, a detection of sources in both the NIR and MIR would be invaluable in studying bolometric outputs, and the effects of internal dust.

What kinds of galaxies will these be? As the universe cools after the Big Bang, the ionized plasma eventually recombines to form a neutral H/He gas at $z \sim 1000$, and photons from this last scattering surface (LSS) form the Cosmic Microwave Background (CMB). However, this neutral gas must have been reionized, as observations of quasars out to $z \sim 5$ reveal no evidence of absorption shortward of rest frame Ly- α by intervening neutral intergalactic H (the Gunn-Peterson effect). On the other hand, reionization must have occurred at $z < 40$, as otherwise photons from the LSS would scatter off free electrons, erasing the observed CMB anisotropies. The most likely explanation for reionization is photoionization by the first generation of galaxies and quasars. Popular Cold Dark Matter (CDM) cosmologies predict the collapse of the first baryonic objects as early as $z \sim 30$ on a mass-scale just above the cosmological Jeans mass, $\sim 10^5 M_\odot$. These

Together, the fine structure and PAH (next section) features offer valuable tools with which to understand the energy generation mechanisms in dusty galaxies, as recently demonstrated by ISO observations of a small sample of the nearest ULIRGs (Genzel *et al.* 1998, Lutz *et al.* 1999). However, even ISO had great difficulty detecting fine-structure lines (equivalent widths of $1 - 4 \times 10^{-2} \mu\text{m}$) in systems at $z \sim 0.1$, requiring many hours of integration time per source. SIRTf will be able to detect these lines in ULIRGs out to $z \sim 1$, but only NGST will be able to obtain spectra of the far more numerous dusty galaxies with $L_{\text{IR}} \sim 10^{11} L_{\odot}$ at these redshifts, and to obtain spectra at even higher redshifts.

[illegible]

environments surrounding an active nucleus (Fig. 8). The low resolution mode ($R = 100$) of an MIR spectrometer on NGST would be well suited to measuring these features to very low fluxes (high- z 's).



The broad dust (and narrow atomic) features can also be used to measure redshifts for obscured, dusty systems that cannot be studied effectively in the optical. Since the

PAH features are very broad and can be strong in starburst galaxies (Fig. 8), they may also function as photometric redshift indicators. This was intended to be the fundamental method used to select moderate redshift ULIRGs in the WIRE database, and will be one way the MIPS instrument on SIRTf will find such systems at high- z . However, it is very important to realize that while the SIRTf cameras will be able to find ultraluminous and “hyperluminous” ($L_{\text{bol}} > 10^{13} L_{\odot}$) galaxies at $z \sim 1 - 5$ and maybe even beyond (the 5σ point source sensitivity of MIPS at $25\mu\text{m}$ will be ~ 0.1 mJy in 1 hr), only NGST will have the sensitivity to obtain spectra of these sources, and perform a study of their energetics and evolutionary states. NGST can reach levels that are a factor of 100 deeper, with arrays having 100 times more pixels. At low resolution, NGST will be able to measure the MIR PAH features in ULIRGs to $z \sim 3$, and hyperluminous infrared galaxies ($L > 10^{13} L_{\odot}$) at redshifts of $z = 5 - 10$ using the $3.3\mu\text{m}$ feature. In fact, the projected sensitivity of a low resolution NGST spectrograph operating at $25\mu\text{m}$ is well matched to the limits of the deep imaging surveys intended for MIPS on SIRTf.

Atomic H Lines. Although more susceptible to extinction than MIR features, NIR lines still offer an order of magnitude decrease in the effects of extinction as compared to UV and optical lines. Strong hydrogen recombination lines such as the Paschen ($n_{\text{lower}} = 3$) and Brackett ($n_{\text{lower}} = 4$) series provide a direct measure of the ionizing flux, and thus the star formation rate (SFR), in galaxies over a wide range of redshift. The NIR lines fill in the critical redshift range of $z = 1 - 3$, where it has been suggested that the overall SFR in the Universe levels off or declines. The recombination line flux ratios also provides a direct measure of the extinction to the energy source(s), and so a route to the true galaxy luminosity. Potentially even more interesting, the strong optical $\text{H}\alpha$ line also enters the MIR for $z > 6.6$. Because $\text{Ly-}\alpha$ is absorbed by the neutral IGM, the $\text{Ly-}\alpha$ line will be unavailable for observing galaxies prior to reionization. On the other hand, Balmer photons are unaffected by large amounts of neutral H, and these lines are thus of paramount importance in finding galaxies beyond the epoch of reionization, and in obtaining redshifts. In addition, since the Balmer lines are less affected by dust than is $\text{Ly-}\alpha$, they also provide a cleaner probe of ionization and star formation rates.

Fig. 9 shows the number of potential sources available in a $2'$ field as a function of z (Oh 1999), in the high-redshift semi-analytic models of Haiman & Loeb (1998). Star formation efficiencies are adjusted to reproduce the observed metallicities of the $\text{Ly-}\alpha$ forest at $z = 3$ (*i.e.*, ~ 1.7 to 17% of the galaxy gas mass fragments into stars),

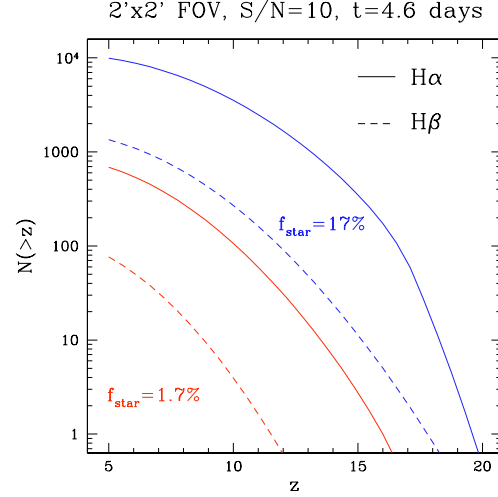


Figure 9: Number of potential H- α sources per FOV. The plot gives to the number of candidate galaxies in a 2-d field, not the number falling on a spectrograph slit.

under the assumption that these systems are pre-enriched by the nucleosynthesis products of the high- z stars. In these models, with an integration time of 4×10^5 sec, the H- α line may be detectable in galaxies out to z in the 5–10 range (Fig. 9).

In addition, because H lines are permitted transitions, they can arise in very dense regions near the galactic nucleus, marking the location of an AGN with the large line widths ($> 10^3 - 4$ km s $^{-1}$) indicative of gas moving in the potential well of a massive black hole. Since H lines are generally narrow in starburst galaxies without a central AGN ($100 - 300$ km s $^{-1}$), high resolution spectra ($R \geq 1000$) can distinguish normal galactic rotation from motions near a massive black hole.

Molecular H₂ lines. The rest-frame NIR also contains sensitive probes of other phases of the ISM. A large number of H₂ lines from $1.5\mu\text{m}$ to $2.3\mu\text{m}$ allow for studies of the warm molecular gas that is ubiquitous in luminous IR galaxies. Nearby IR galaxies contain large amounts of molecular gas (traced in its cold, dense state via mm transitions of CO) which, because it is highly dissipative, settles in the evolving potential well of merging galaxies and produces stars at fantastic rates ($10 - 100 M_{\odot} \text{yr}^{-1}$). Measurement of the H₂ lines would allow an estimate of the mass of warm molecular gas in high- z galaxies, and a determination, via the line ratios of rotationally and vibrationally excited features, of the heating mechanism(s). The line ratios can be used to differentiate between excitation via X-rays or UV photons, or the action of fast shock waves.

Fe lines. The NIR is also home to six [FeII] lines in the $1.25\mu\text{m} - 3\mu\text{m}$ region, which function as tracers of the energetics of the ISM. While most of the Fe in quiescent galaxies is locked up in grains, making these lines relatively weak, in galaxies with rapid star formation, the NIR [FeII] lines can be very strong because shocks (from supernovae, large-scale winds, or galaxy collisions) destroy the grains and heat the Fe-rich gas. Thus, strong [FeII] emission is an excellent tracer of very energetic activity in galaxies. In a sample of dusty, possibly merging galaxies at high- z , the [FeII] lines will pinpoint those sources with the largest supernovae rates and the strongest outflowing winds. Used in concert, the H_2 and the [FeII] lines will provide much needed insight into the processes which regulate star formation in galaxies at high redshift and the methods by which galaxies transport energy and chemically enriched material into the intergalactic medium.

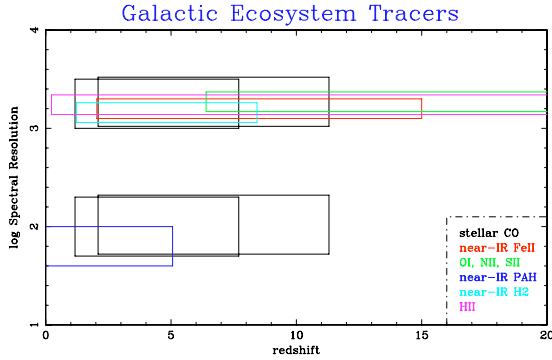


Figure 10: Desired instrumental resolving power for the various spectral features discussed. The colors indicate the following features – black: stellar CO, red: NIR [FeII] lines, green: low ionization optical lines of [OI], [NII], and [SII], blue: NIR PAH features, cyan: NIR H_2 lines, magenta: H-recombination lines

Spectroscopic Requirements. All of the spectral features discussed can be used in concert to assemble a coherent picture of “galactic ecosystems” at early epochs, when starbursts and galactic mergers were creating the range of disk and elliptical systems we see today. To meet these goals, a combination of low ($R \sim 50 - 100$) and high ($R \sim 1000 - 3000$) resolution spectroscopy covering the full $5 - 30\mu\text{m}$ region is called for (Fig. 10). Through careful measurement of the MIR and rest-frame NIR emission features from atomic and molecular gas and dust, NGST would then be able to delineate the star-formation process at the earliest epochs, which led eventually to the variety of structures seen along the Hubble sequence today.

2.3 Stellar Populations and Galaxy Morphology

An MIR-capable NGST also has a critical role to play in the measurement of young and old stellar populations in galaxies, and in the definition of galaxy morphologies. Most galaxies are made up of a mix of stellar populations, the fraction of young and old stars changing both with redshift (age) and Hubble type. The young stars can be measured directly in the UV (if unobscured), or inferred spectroscopically from their effects on the surrounding ISM (as in Sect. 2.2). To see the old stars, however, it is crucial to sample the rest-frame NIR spectrum of a galaxy, where the 1.62 and $2.29\mu\text{m}$ CO bandheads lie, and this can only be achieved in the MIR for galaxies at cosmologically interesting redshifts. Such measurements are very important for studies of the evolution of many different types of galactic systems: In galaxies selected to be young (*e.g.*, Lyman break galaxies) one can search for old stars that would signal a previous major episode of star formation. In systems known to be old, or where a massive galaxy is likely already in place (*e.g.*, high redshift radio galaxies and/or QSO hosts) one can search for evidence of significant young stellar populations (also by means of the CO bandheads) that may have been triggered by the passage of radio jets or the merger of galaxies.

Imaging. At this point, the Lyman break technique has generated hundreds of galaxies at redshifts of $z = 2.5 - 4.0$. As this technique is pushed to longer and longer wavelengths over the next few years (*e.g.*, B, R, and I-band dropouts), the numbers of known galaxies in the $z = 4 - 7$ range will greatly increase. Presently, the break technique naturally selects for UV-bright, blue galaxies, and therefore is very good at finding rapidly star-forming systems (with some AGN sprinkled in as well). Since the redshifts of these galaxies are close to those during which large galaxies were still being assembled, it is natural to ask whether or not the break galaxies are going through their first episodes of star formation, or whether we are seeing large bursts on top of pre-existing stellar populations. The former would be considered as more “primeval galaxies” while the latter might be more like ULIRGs in terms of their star formation rates (although probably not as dusty). Therefore, it is critical to observe these sources at wavelengths well beyond those used to find them in the first place (the rest frame UV) because that is where the old(er) populations will dominate the integrated light. The best place to look for old stellar populations is in the rest frame H or K bands. In Fig. 11 we show the WFPC and NICMOS images of a faint galaxy in the Hubble deep field (Thompson *et al.* 1999). The rest-frame optical emission clearly gives a very biased and incomplete picture

of this galaxy: it completely misses the central, bulge-like feature so prominent in the NIR. The resulting morphological disparity can be dramatic at high- z , making an MIR capability critically important.

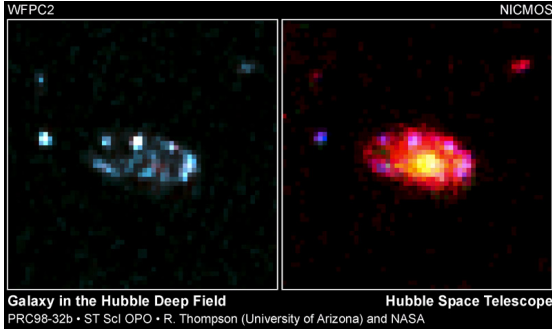


Figure 11: Rest-frame B and NIR images of a faint galaxy in the HDF, highlighting the very different stellar populations traced at each wavelength. Image courtesy NASA/STScI.

While the Lyman-break galaxies are a good example, the same arguments apply to any galaxies at $z > 2 - 3$, even the host galaxies of AGN (*e.g.*, QSO's and radio galaxies). It will be easier to detect the host galaxies and measure the old stellar populations (if they exist) in these presumably massive systems by using a high-resolution MIR camera on the NGST. The old stars should not only be present, but they may also have a “relaxed” stellar distribution, similar to an elliptical galaxy today. A luminous old stellar population with a relaxed stellar light distribution at high- z makes a strong argument for a massive galaxy already in place at high- z , and this, in turn, places very strong constraints on the cosmology within which these systems were assembled. For example, Benitez *et al.* (1999) report the detection of a substantial population of $1 < z < 2$ early-type galaxies with old (≥ 2 Gyr) stellar populations, based on recent NICMOS & VLT observations of the HDF-South field.

Even if old, fully-assembled galaxies at $z > 3 - 5$ are not found, it still might be possible to detect sub-galactic building blocks which formed first, *i.e.*, bulges, and constrain their formation epoch. As discussed above, the best way to find aging bulges at significant z is in the MIR, where the old stars peak and the extinction is at a minimum. Local bulges have K-band surface brightnesses of $16 \text{ mag arcsec}^{-2}$, corresponding to $250 \mu\text{Jy arcsec}^{-2}$. At $z=4$, given the cosmological dimming factor, $(1+z)^3$, the surface brightness of these bulges will be $\sim 2 \mu\text{Jy arcsec}^{-2}$ or $0.08 \mu\text{Jy/beam}$, shifted to the $10 \mu\text{m}$ band. This $10 \mu\text{m}$ surface brightness should be detectable ($S/N \sim 10$) in about 16 hours, even with no evolution (brightening) in

the bulge surface brightness. The physical sizes of local bulges (*e.g.*, in M31) are $\sim 4 \text{ kpc}$, which at $z=4$ corresponds to an angular size of $0.5''-1''$. Such bulges can be resolved with NGST, enabling the study of their light distributions, which may provide important (dynamical) information on when and how bulges form.

Finally, as touched upon earlier, the contribution of accretion-powered energy from AGNs to the total energy density of the universe remains an outstanding unsolved problem, as are the related issues of the first generation of AGNs and the co-evolution of massive black holes and galaxies. Star formation is most often the source of the radiation that heats the warm dust in galaxies, but in active systems, an AGN can heat dust to very warm temperatures and generate blackbody spectra which peak in the MIR. In cases where most of the energy is produced in a small volume (most active galaxies as well as ULIRGs) it is important to have the highest possible spatial resolution to correctly assign the relative fraction of energy produced by hot stars and an AGN. In section 2.2 the means by which MIR spectroscopy can unearth buried power sources was described, but high resolution MIR imaging has an important role to play here as well (*e.g.*, Fig. 12). Since an AGN is basically unresolved, while a starburst will extend from several hundred pc to a few kpc in diameter, the size of the warm dust region in the nucleus of an active galaxy can directly constrain the energy source. In this manner, an MIR imager becomes a very powerful tool for probing the central regions of dusty galaxies, and for determining whether the energy is deposited by central accretion or by young, hot stars.

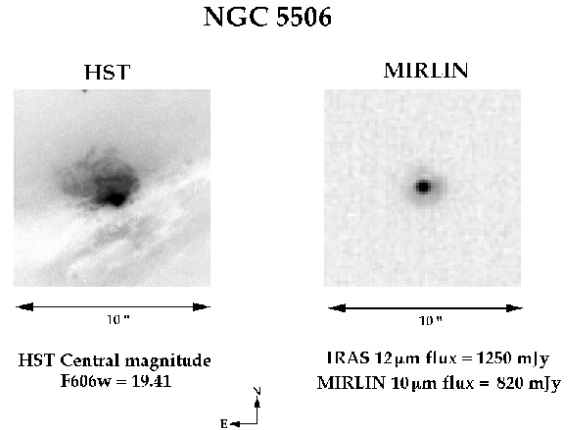


Figure 12: NGC 5506 is a spectacular example of an optically invisible AGN being revealed in the MIR. The image on the left is from HST, and the right-hand image from the MIR camera MIRLIN. Note the Airy diffraction ring on the unresolved AGN in the MIR. (Courtesy of V. Gorjian)

Spectroscopy Although deep MIR imaging at high spatial resolution is the easiest way to obtain the initial evidence for an old population of stars in a high redshift galaxy, the ability to measure stellar absorption lines places the study of early galaxy formation and evolution on a much more quantitative footing. Stellar photospheric features in the rest-frame optical and NIR produce strong absorption in the I, H and K-band spectra of galaxies. These features can be used to infer the age and metallicity of the stellar population, as well as the total dynamical mass of the high redshift galaxy. The calcium IR triplet (8498, 8542, 8662 Å), visible at $z \geq 4.8$, is an excellent metallicity indicator for stellar systems with ages ≥ 1 Gyr. The CO bandheads, on the other hand, can be exploited for their sensitivity to the effective temperature of the stars. At even higher redshifts ($z \geq 9$), traditional optical spectral indices such as H β , MgII 5175 Å, and Fe 5270 Å can also be used.

Since young starburst galaxies (at ages of 1–few $\times 10^7$ yrs) can have unusually strong CO absorption bands due to the presence of large numbers of red supergiant stars, the simple comparison of the CO band strengths with other diagnostics (*e.g.*, the adjacent Br- γ line equivalent width, or the rest-frame K-band luminosity) can be used to estimate the most important parameters of the burst (age, mass fraction involved, etc. (Armus *et al.* 1995). The easiest way to gauge the overall strength of the CO lines is with a low resolution ($R \sim 100$) MIR spectrograph. In Arp-220-like galaxies, for example, the CO bandhead can be detected out to $z = 4$ in several hours, and out to $z \approx 5$ in 1 day, thus allowing a search for strong CO absorption in the integrated light of galaxies during the time when massive galaxies were first being assembled and star formation was at its peak.

At higher resolution, a very important application of the stellar absorption-line data is to constrain the dynamical mass of the systems. To date, this has been one of the key missing links that prevents a straightforward comparison between observations and theoretical simulations of structure formation. With this goal in mind, it is imperative that the spectra have sufficient S/N and spectral resolution to measure the velocity dispersions of the stars. The calcium IR triplet lines are ideal for this application. Because little is known about the mass distribution of galaxies at high redshift, a resolution of $R \sim 3000$ would be desirable to ensure that late-type galaxies and protogalactic clumps with internal velocity dispersions as low as ~ 20 – 40 km s $^{-1}$ can be detected. The NIR CO features can also be effectively used for this purpose, since they are commonly the strongest stellar absorption features in the H and K bands of normal and starburst galaxies. High resolution

spectroscopy of the CO lines, when coupled with observations of the galactic light profile and simple models of the stellar distribution, can be used to estimate the mass of the galaxy (Shier, Rieke & Rieke 1994). By comparing the dynamical mass derived from either the Ca or CO absorption lines to that of the cold gas mass derived from millimeter CO emission-line observations, an estimate of the evolutionary state of the galaxy can be made.

2.4 The Lifecycle of Interstellar Grains

Interstellar matter provides the basic building blocks from which new solar systems like our own are made. The formation of stars begins with the collapse of a dense interstellar cloud core, a reservoir of dust and gas from which the protostar and circumstellar disk are assembled. The dust particles are typically 0.01–0.1 μ m in size, and consist of silicates, oxides and carbonaceous material such as PAHs, formed in the outflows of dying stars. In cold environments, gas-phase molecules condense on the grains and form an icy layer. During the evolution from their formation sites in aging stellar envelopes through their stay in the ISM until their incorporation into new planetary systems, these solid-state species undergo a complex series of physical and chemical metamorphoses. Solid-state vibration-band spectroscopy in the MIR (3–30 μ m and longer) is the *only* way to probe this dust lifecycle (Fig. 13).

Generation of Dust. In the last phases of their lives, stars of moderate mass begin to pulsate and lose their outer layers. The material ejected during this stellar death struggle is enriched in heavy elements which have been formed by nucleosynthesis inside stars. In the cooling outflow, a complex chemistry develops which allows solid particles to condense. If the outflow is O-rich, silicates and oxides are created (Fig. 14), whereas in C-rich environments carbonaceous material and PAHs are formed (Fig. 15).

However, thermodynamic equilibrium is not always obeyed, and kinetic effects can play a role. PAHs are thought to be formed in the pulsation-driven shocks and gas excursions close to the stellar photosphere through a combustion-type chemistry. Recent ISO spectra (Figs. 14 & 15) at 5–45 μ m indicate a rich variety of amorphous and crystalline minerals and PAH features in late-type stars. CRL618 in particular has been found to be a veritable factory of exotic carbon-rich species (Fig. 15).

These variations, as well as the origin of the crystalline material, are not well understood at present. NGST can provide major advances in this area by providing spatially resolved imaging of these features in nearby objects, and by probing the evolution of the solid-state features in more

Solid State Features	Wavelengths (μm)
PAHs	3.3, 6.2, 7.7, 8.6, 11.0, 11.2, 12.7, 15.9
Amorphous silicates	9.7, 18
Crystalline silicates & oxides	
Mg_2SiO_4 (Forsterite)	10.2, 11.4, 16.5, 19.8, 23.8, 27.9
FeO	3, 23
Ices	
H_2O	3, 6.0
CH_3OH + uniden	3.5, 6.8
CO_2	4.27, 15.2
XCN	4.62
CH_4	4.67
NH_3	9.7
Molecular lines and bands	
H_2	28.21 S(0), 17.03, 12.28, 9.66, 8.03, 6.91 S(5)
CH_4	7.7
C_2H_2	13.7
HCN	14
CO_2	14.98

Figure 13: MIR Dust Features and Molecular Lines

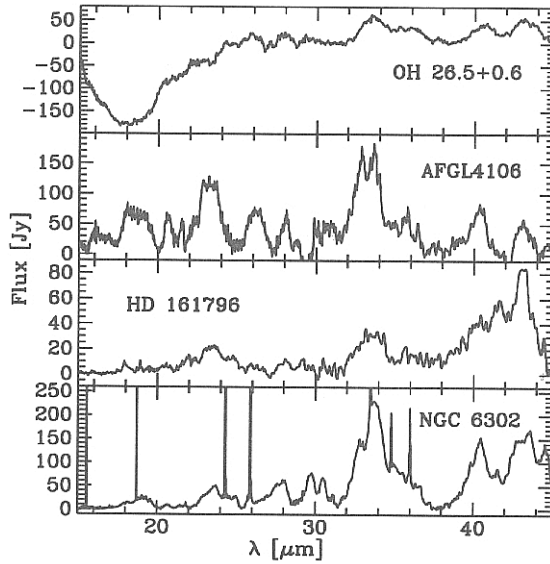


Figure 14: ISO SWS spectra of the AGB star OH 26.5+0.6, the red supergiant AFGL4106, the protoplanetary HD161796, and the planetary nebula NGC 6302 (Tielens *et al.* 1998), illustrating the richness and variety of MIR circumstellar silicate spectra.

distant Galactic objects and galaxies with different initial elemental abundances. Once ejected, the composition of carbonaceous dust is affected by exposure to intense UV radiation, such as is experienced during the planetary nebula phase and subsequently in the diffuse ISM. Generally, the radiation ionizes PAHs and removes H from the carbon skeleton, leading to variations of more than a factor of 5 in the ratios of the 8.6/7.7 and 11.3/7.7 μm features, and significant other changes to the 11–20 μm spectra.

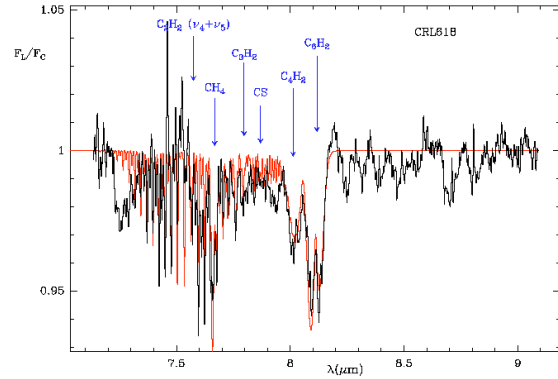


Figure 15: ISO spectrum of CRL618. Many bands are identified with new molecules such as C_4H_2 and C_6H_2 ; many others remain to be identified. (Courtesy J. Cernicharo & J. Pardo)

Incorporation into molecular clouds In the cold and dense phase of the ISM, gas-phase molecules freeze-out onto the grains and form an icy mantle surrounding the silicate and carbonaceous cores. In the coldest cores, the ices can contain the majority (>50%) of the condensible elements (*i.e.*, C and O), and can be studied only through absorption spectroscopy against background infrared sources. Most such studies to date have been performed toward deeply embedded massive young stellar objects (YSOs) and a few bright field stars that happen to lie behind dark clouds. These field stars are the best probes of the dust in quiescent cold clouds, where no heating or processing has yet taken place. Although SIRTf will further such studies, the unprecedented MIR sensitivity of NGST can revolutionize this field by making such studies possible toward hundreds of background stars in a large variety of clouds with different physical conditions, thermal histories, and dust columns. Moreover, the dust in the earliest, deeply embedded collapse stage of star formation, where depletion should be at its maximum (>99% of the heavy elements frozen out), can be studied using the newly discovered Class 0 protostars as targets, primarily at the longer MIR wavelengths, because of their plunging MIR spectra (Fig. 20).

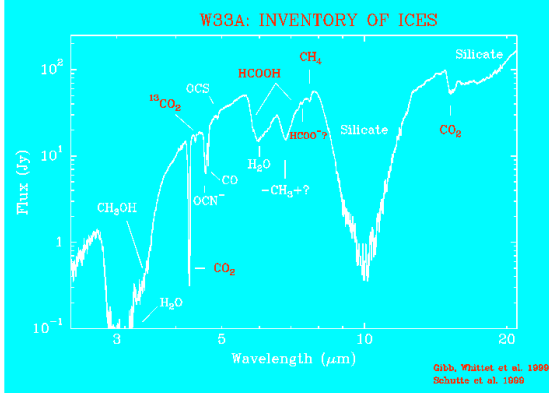


Figure 16: ISO spectrum of W33 showing wealth of ice features (Gibb *et al.* 1998; Schutte *et al.* 1998).

Once a young star has formed, its radiation will heat the ices and process the material. Most of our information to date stems from ISO observations of only a dozen very luminous high mass young stellar objects (YSOs), whose spectra show a wealth of ice features, including H_2O , CO , CO_2 , CH_4 and CH_3OH , arising in the dense envelope around the star (Fig. 16). NGST, with a factor of 1000 greater sensitivity, will open up such studies to young stars of all masses at all evolutionary stages, including objects similar to our early Sun. Many of the features, in particular the CO_2 band at $15\ \mu\text{m}$, are very sensitive to the thermal history of the material, in particular to the cumulative amount of heating that the ice has experienced (Fig. 17). Other features, such as the unidentified band at $4.62\ \mu\text{m}$, may be sensitive primarily to ultraviolet processing. Thus, these features could be suitable recorders for relatively short-lived evolutionary phases of intense radiation, such as FU Orionis events which signal intermittent high disk accretion rates in YSO's.

The heating of the dust also results in evaporation of the ices, returning molecules such as H_2O and CH_3OH back to the gas phase. These species subsequently drive a rapid chemistry in which complex organic molecules are formed. At sufficiently high resolving power (a few thousand), NGST will be able to detect these evaporated molecules directly (Fig. 18). The combination of UV radiation and heating of ices also leads to a more refractory organic component, which will eventually be incorporated into new solar systems.

Circumstellar disks and connection with solar system matter. The solid material in star-forming regions may eventually be incorporated into circumstellar disks and solar system material, including comets and Kuiper-Belt objects. The close connection between material in disks

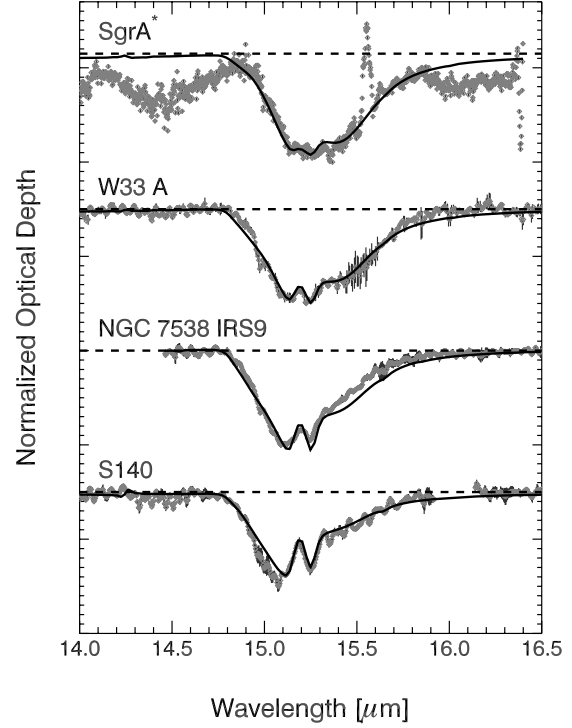


Figure 17: ISO spectra of a prominent bending mode of CO_2 ice, showing the re-structuring and segregation of ices on grains, as a function of increasing temperature from low (top: $T \sim 20\text{K}$) to high (bottom: $T \sim 80\text{K}$) (Langer 1999).

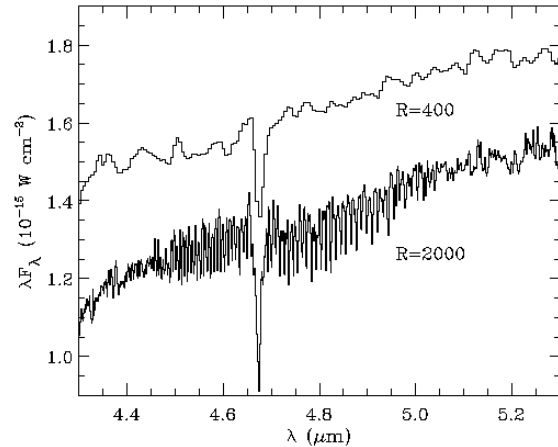


Figure 18: Upper spectrum: low-resolution ($R=400$) ISO spectrum of CO towards a nearby YSO. Lower spectrum: same spectrum at high-resolution ($R=2000$). The CO gas phase is only evident at high resolution (Boogert 1999).

and comets is illustrated in Fig. 19, where a remarkable similarity between the solid state features (PAHs and crystalline silicates) is evident. Interestingly, a fraction of the silicates appears to be in crystalline form in one specific mineral, forsterite (Mg_2SiO_4). The origin of the crystalline material is not yet understood, but must be related to the heating events experienced in the disk. Spectra of other disks show a wide variety of other crystalline and amorphous material.

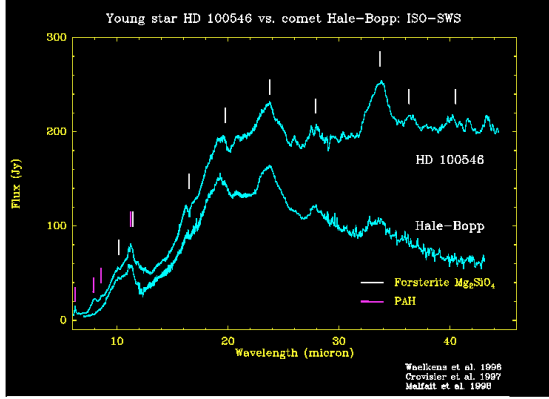


Figure 19: Comparison of cometary and YSO disk spectra (Waelkens *et al.* 1996).

ISO observations have been limited to only a handful of unresolved disks around intermediate mass stars, but NGST will be able to image such disks in various MIR features around lower-mass, solar-type stars, and perform absorption spectroscopy of ices and other features toward the highly-extinguished stars behind edge-on disks, such as those embedded in the proplyd disks in Orion. In summary, only an MIR spectrometer on NGST will be able to cover the relevant wavelength range, and to provide the sensitivity and spatial and spectral resolutions necessary to probe the full lifecycle of dust from its formation in dying stars to its incorporation into new planetary systems.

2.5 Protostars

A protostar is an object which derives its luminosity from the gravitational collapse of a dense, pre-stellar gas and dust cloud core. Understanding how these objects form and evolve is key to understanding the formation of single and binary stars and their accretion disks, which eventually give birth to extrasolar planetary systems. A decade ago, protostars were still considered the “Holy Grail of Infrared Astronomy” since direct NIR emission from a protostar had never been found. Instead, these rare, short-lived, and extremely obscured objects were first recog-

nized at submillimeter wavelengths, because the dust in an infalling protostellar envelope (of typical dimension a few thousand AU) reaches very high extinctions ($A_V > 400$), effectively cutting off the spectra in the MIR (Fig. 20).

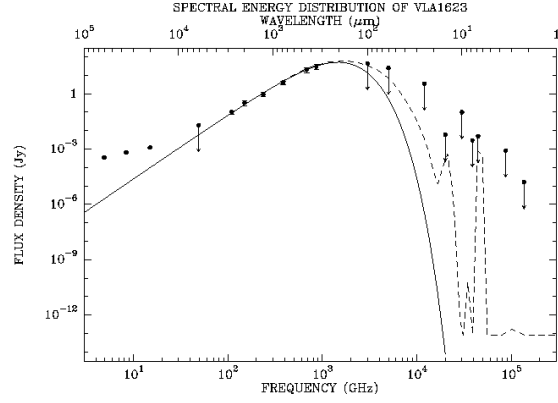


Figure 20: Spectrum of the prototypical Class 0 protostar VLA1623. Such highly extinguished spectra require 20 – 30 μm observations for progress. (Courtesy P. André).

The youngest (few $\times 10^4$ yrs) identifiable protostars all drive powerful, bipolar gas jets, which abate with age, disappearing by the time the central object has attained its initial stellar mass. The origin of these bipolar outflows is not yet well-understood, but is clearly intimately related to the infall process and the transport of excess energy and angular momentum away from the central object. In fact, it is precisely the presence of the bipolar cavities excavated by protostellar outflows which will allow the direct detection and imaging of the thermal dust emission from protostars at MIR wavelengths. To date, all 45 known protostars within 500 pc of the Sun remain either undetected in the MIR (Fig. 20), or, if detected by ISO, appear as weak (mJy level) point sources. With an MIR capability reaching to 30 μm , NGST can revolutionize our knowledge of the star formation process, by providing us with our first detailed look at the highly obscured, warm (few 100K) inner dust structures closest to the central accreting object (Figs. 21 & 22).

An imaging capability to $\lambda \approx 30 \mu\text{m}$ is crucial for discerning the infalling envelope structure of the youngest protostars, as demonstrated by the model depicted in Fig. 21. This figure shows the appearance of a nearly edge-on infalling protostellar source, with a 10AU radius accretion disk, as it would appear to NGST’s MIR camera at 10, 20, and 28 μm . The assumed infall rate is $2 \times 10^{-5} M_\odot/\text{year}$, the total envelope extent 3000 AU, and the cone-shaped bipolar cavity opening angle 10° . The visual extinction through the envelope, $A_V = 590$, depends on the central

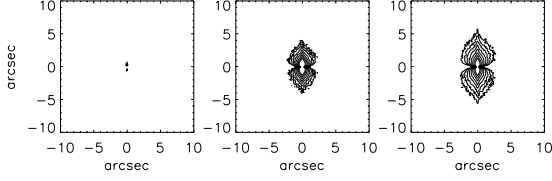


Figure 21: Simulation of the structure of VLA1623, an almost edge-on protostar 125 pc away, at 10, 20, and 28 μm (l to r). The integrated fluxes are 0.01, 2.6, and 1.4 mJy (for standard interstellar dust grains (with $R \equiv A_V/E(B-V) = 3.1$) and an infall rate of $2 \times 10^{-5} M_{\odot}/\text{yr}$).

object’s mass, accretion disk size, and the infall rate, a quantity that has thus far eluded direct measurement. Interestingly, the nascent flared disk of 10 AU radial extent, although completely unresolved at the assumed distance of 125 pc, nevertheless projects a huge shadow, which far exceeds the physical size of the disk. Images such as these could be obtained in just a few hours with NGST, whereas they are impossible to obtain from the ground.

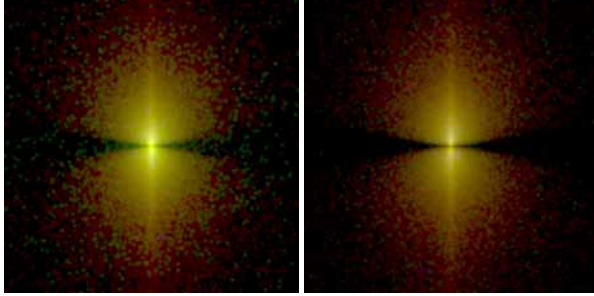


Figure 22: False-color MIR images of the infall model of Fig. 21, with 10/20/28 μm emission in blue/green/red. The redder image has $R = 5.5$ molecular cloud grains, the bluer one standard $R=3.1$ interstellar grains.

Multi-color imaging will also be able to discern the evolution of dust from interstellar to protostellar grains, since the MIR colors of protostellar infall envelopes will be greatly affected by the dust grain properties. For example, in Fig. 22 we show in false-color the same infall model as in Fig. 21, but for two different types of dust grains: “standard” ($R=3.1$) interstellar grains seen in the diffuse ISM (left) and the larger grains typical of dense, dark clouds such as Ophiuchus (right). The larger grain model is clearly “redder”. Since higher extinction will also have some reddening effect, multi-wavelength (including mm/submm) observations will likely be needed to disentangle the effects of grain size distributions and grain types from grain column densities.

NGST studies of protostellar envelopes will be complementary to the information garnered with the ALMA. In recent OVRO studies of protostars in Serpens, for example, it was found that the central continuum sources in these objects, where detected, corresponds to the envelope regions rather than to unresolved disks, and that a warm (few $\times 100\text{K}$) component, distributed over larger scales than the inner few hundred AU, is required to account for the 1.4mm continuum flux (Hogerheijde *et al.* 1999). Such a warm, extended dust component should be readily observable with an MIR imager on NGST, and the state of the gas explored via MIR spectroscopy.

The relative strengths of various emission lines (especially of molecular H_2 in the 6–28 μm range) indicate the shock velocities of the material impacting the disk, which provides an indication of the disk impact radius. Together with the observed bolometric luminosity, this leaves only the product of the infall rate and the central object mass as an unknown. However, quantitative modeling of the MIR appearance of a protostar at four or more wavelengths (as above) can sufficiently narrow the range of allowable infall rates and central source masses to allow the required order-of-magnitude improvement in measuring the elusive infall rate, whose value will distinguish amongst competing star-formation theories.

MIR imaging and spectroscopy with NGST can thus make fundamental contributions to many outstanding problems in star-formation, including the determination of infall/outflow geometries, the estimation of infall rates, the measurement of temperature, density, and grain-property distributions in the infall region, and the elucidation of formation mechanisms (single vs. binary systems).

2.6 Lifecycles of Disks

Eventually, adolescent stars emerge from their cloudy wombs, allowing more direct observation of disk structures. These young disks lead to planet formation, and so it is vital to study such disks over a range of ages. Spectral models suggest that as young T Tauri (TT) stars begin to lose the $\text{H}\alpha$ signature of disk accretion (*i.e.*, progress from a “classical” to a “weak-line” TT phase), small dust grains initially disappear from the inner disk. It is likely that as molecular gas disappears, viscous accretion fails to replenish grains which spiral into the star. It may also be the case that the grains have aggregated into planetesimals. As this process continues, it becomes impossible to detect residual IR excesses from the ground. It may be that these stars have particulate disks like that around β Pic, but at the distances of the nearest star-forming regions, and at the lower temperatures commensurate with

the environments of low-mass stars, they are far more difficult to detect. MIR imaging on NGST can excel at detecting the faint infrared excess from stars at this phase and at resolving structure. This will fill a much-needed gap in our understanding of the post-classical TT evolution of circumstellar disks. In addition, it will be possible to image holes and/or gaps expected in disks in which planets form..

At ages of 10^7 – 10^8 yr, circumstellar material was detected by IRAS around nearby (10–100 pc) A stars in the thermal IR. A stars are brighter than low-mass stars, and can heat material to temperatures corresponding to MIR wavelengths at larger radii from the star. IR images and spectra of HR 4796 (cover), and submillimeter images of Vega, Fomalhaut, β Pic, and ϵ Eridani (Fig. 23) reveal that, at this phase the disks possess inner solar-system-sized regions in which the density of grains is vastly reduced. These disks are ephemeral; they have persistence timescales that are much smaller than the age of the system. It is theorized that they are maintained by collisional action within a system of planets and planetesimals. Observations of HR 4796A, for example, reveal a radial dust distribution that resembles an enhanced version of our own solar system's zodiacal dust and Kuiper Belt distributions. The presence of planetary bodies embedded in these disks is expected to be evident in the small scale morphology of the dust. Azimuthal asymmetry is already suggested in the $15''$ resolution JCMT/SCUBA images. High-resolution NGST/MIR observations (at $0.25''$ resolution) will be especially capable of detecting these signatures. For disks like that around Vega, the surface brightness of the dust is too low to be detectable in the MIR from the ground, but NGST/MIR observations will be able both to detect and resolve AU-scale details.

Disks in the 10^8 to several $\times 10^9$ year age range are the evolutionary bridge between planet-forming disks and remnant disks like our solar system. NGST MIR imaging and spectroscopy of warm dust around main sequence AFGK stars within 10 pc (combined with ancillary effort to determine stellar ages) can place our system in an evolutionary context, and potentially quickly reveal the dynamical influence of planets. The early part of this span (1 to 5×10^8 yr) corresponds to our heavy bombardment era, during which the terrestrial zone of the planetary system would have had a dust density a few thousand times the present value, presenting a strong MIR signal from collision debris. Assuming a flat age histogram, $\approx 10\%$ of nearby stars should fall in the heavy bombardment age range. There is a controversy about whether this heavy bombardment was simply the tail of the planet accretion process or a discrete event caused by the final construc-

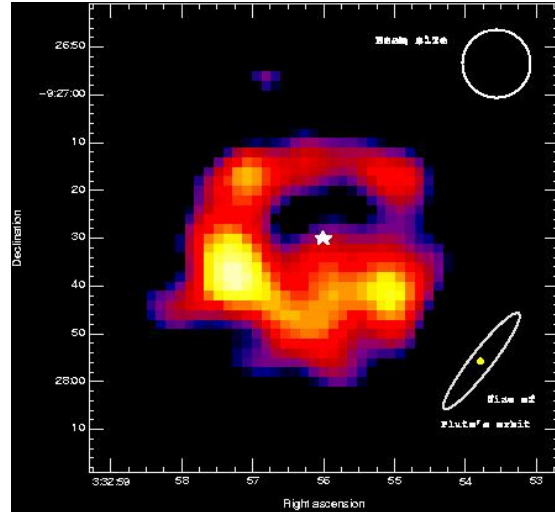


Figure 23: JCMT SCUBA $850 \mu\text{m}$ image of the ϵ Eridani debris disk. (Greaves *et al.* 1998).

tion of the outer planets and the scattering of planetesimals across the solar system. It is probably significant that life did not begin on earth until the bombardment ended. Heavy bombardment is also a possible means of introducing much of Earth's biospheric organics. Finally, low upper limits on warm (terrestrial-temperature) dust can indicate systems toward which the proposed Terrestrial Planet Finder (TPF) interferometer could preferentially detect faint flux from terrestrial planets.

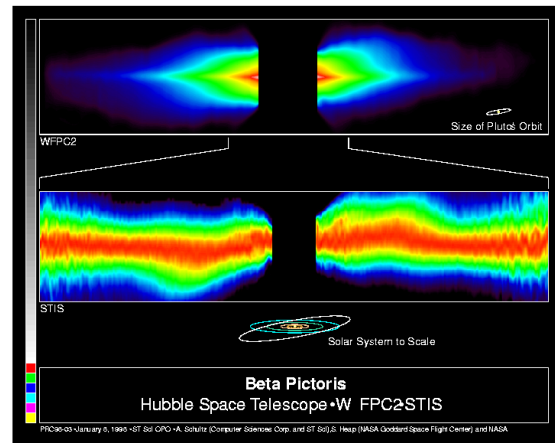


Figure 24: STIS image of the warp in β Pic's disk, a tell-tale of a planet. (Courtesy STScI.)

At the present age of 5×10^9 yr, all the dust in our solar system is necessarily material recently released by collisions and sublimation of remnant planetesimals (as-

teroids and comets). An external observer of our system could infer the presence of planetary-mass objects both as the perturbers of asteroids and comets into crossing and star-grazing orbits, and as the cause of warps, asymmetries, density waves, empty annuli, and sharp edges in the populations of dust and dust parent bodies. One example is the resonant ring and wake of zodiacal dust trapped in the Earth's orbit (Dermott *et al.* 1994), seen by IRAs and COBE. A second example (Fig. 24) shows the mid-plane warp in the β Pic disk, modeled as being due to a Jovian-mass planet. Also relevant in this context is the study of our own solar system's Kuiper belt. NGST can make a contribution here, but the extremely cold temperatures in the Kuiper belt make SIRTf more of a natural match.

Disappearance of Molecules The bulk of the material residing in accretion and remnant gaseous disks is in the form of molecular H_2 . The fundamental rotational transition of H_2 lies at $28\ \mu\text{m}$, and observations of this line provide direct information on the amount of warm molecular gas in star-forming regions and young disks at temperatures of $\sim 60\ \text{K}$ or higher (Fig. 25). Subsequent transitions occur at $17\ \mu\text{m}$, $12\ \mu\text{m}$ etc. (Fig. 13), and are more sensitive to warmer gas at $\sim 1000\ \text{K}$, such as is found in gas heated by ultraviolet photons. A key program for NGST can be the search for and imaging of the low-lying H_2 lines in circumstellar disks to constrain the gas/dust ratio and so the gas dissipation time scale. Observational constraints on this timescale are essential to test theories of giant planet formation. NGST will be sensitive to gas masses of $10^{-5}\ M_\odot$ at $T > 100\ \text{K}$, and $10^{-3}\ M_\odot$ at $T > 50\ \text{K}$, and with multiple-line data, the disk temperature structure can be mapped.

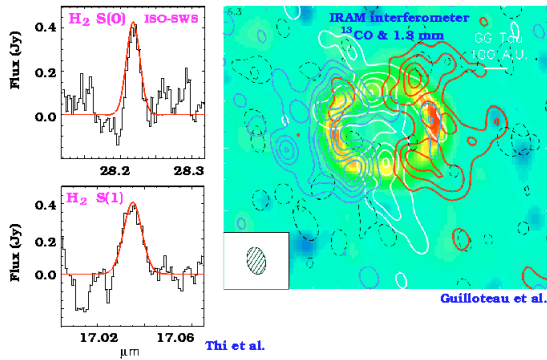


Figure 25: H_2 lines detected by ISO in the GG Tau disk (Thi *et al.* 1999; Guilloteau *et al.* 1999).

2.7 Brown Dwarfs and Extrasolar Giant Planets

Sub-stellar objects (Brown Dwarfs or BD's) with masses ($10M_J$ – $70M_J$) too small to support stable H-burning have been found in three environments: as companions to more massive stars, in a high-luminosity phase in forming star clusters, and as field objects in sky surveys such as 2MASS, Sloan, and DENIS. The field objects have been classified into the L dwarf and T dwarf spectral types (Kirkpatrick *et al.* 1999, Burgasser *et al.* 1999) with effective temperatures in the range 950 – $1500\ \text{K}$ and luminosities $\sim 10^{-4}\ L_\odot$. While masses are hard to determine because of rapid cooling, comparisons with spectral models (Burrows *et al.* 1997) suggest ages of ~ 0.5 – $5\ \text{Gyr}$ and masses of 20 – $70\ M_J$. Objects of still lower mass, 0.5 – $20\ M_J$, have been found orbiting nearby stars via radial velocity oscillations. What is the genesis of these objects? Do brown dwarfs form via the collapse of cold, dense cores that also leads to hydrogen-burning stars? Do massive planets form via agglomeration of material in a protostellar disk? How do these processes overlap? Is there a crossover mass at which one formation process becomes dominant? What is the physics that controls the appearance and evolution of these objects? These are all questions that MIR observations with NGST can address. A MIR program might consist of:

- A modest area sky survey ($10\ \text{sq. deg.}$) at 5 – $10\ \mu\text{m}$ to determine the mass function of the coolest objects in the field down to a level of 1 – $5\ M_J$.
- Directed surveys at 5 – $10\ \mu\text{m}$ of nearby star-forming regions, *e.g.*, TW Hydrae and Taurus, to look for objects smaller than $1\ M_J$.
- Searches for Jupiter-mass companions of nearby stars at 5 – $10\ \mu\text{m}$.
- Spectroscopic observations to determine the physical state of these objects.

Sub-stellar Mass Function. An MIR NGST could undertake 3 surveys to investigate the distribution and evolution of brown dwarfs. Using the predicted MIR flux densities for BD's of various masses and ages (Burrows *et al.* 1997) and extrapolating from a $dN/dM \sim M^{-1.3}$ power law with $0.1\ \text{objects pc}^{-3}$ in the 75 – $100\ M_J$ range (Reid *et al.* 1999), it is possible to estimate (very roughly) the limiting distances and expected BD counts that would be found in a modest NGST survey. Roughly one week would suffice to survey $10\ \text{sq. deg.}$ to the sensitivity levels given in Table 1, resulting in the detection of thousands of BD candidates with masses as low as $5\ M_J$. Three

wavelengths were selected from the Burrows *et al.* (1997) spectra to include a bright peak at 5 μm , a deep CH₄ absorption band around 6–8 μm , and a continuum band at 10 μm . Sub-stellar objects would stand out very clearly from other stars and extra-galactic objects in color-color plots made using these wavelengths.

		5 μm		8 μm		10 μm	
Sensitivity (5 σ , 120 sec)		0.08 μJy		0.3 μJy		0.6 μJy	
Mass range (M_J)	Age (Gyr)	<u>D</u> max (pc)	<u>N</u> um- ber	<u>D</u> max (pc)	<u>N</u> um- ber	<u>D</u> max (pc)	<u>N</u> um- ber
0.85-1.2	1	110	530	6	0	18	2
	5	10	1	0.2	0	1	0
4.2-5.8	1	200	1850	33	9	75	100
	5	110	330	19	2	20	2
17-23	1	>500	>2x10 ⁴	330	5800	330	5800
	5	440	1.4 x10 ⁴	100	180	130	360

A second survey, for free-floating sub-stellar objects, would be made toward nearby star-forming regions, to pin down the low mass cutoff of the fragmentation and collapse process — after all, the collapsing core does not know about the hydrogen-burning limit! The extra luminosity of a young BD/planet more than compensates for the increased distance to the nearest star-formation regions 50–150 pc away. The brightness of a 1 M_J BD/planet at an age of 0.1 Gyr is such that the sensitivity levels in Table 1 are adequate to detect such objects at 500 pc at 5 μm and at 150 pc at 10 μm ; higher mass objects of the same age are obviously brighter and easier to detect. The characteristic peak at 5 μm and the deep CH₄ absorption at 8 μm would help distinguish BD/planets from background objects.

Finally, NGST could be used to detect Jupiter-sized objects orbiting nearby stars. Gaidos (1998) has identified 38 G0–K2 stars with ages < 1 Gyr within 25 pc. Even ignoring heating by the central star, objects as small as 1 M_J (emitting 10 μJy at 5–10 μm at an age of 1 Gyr) would be detectable outside \sim 5-6 Airy radii of the star (corresponding to 8–20 AU orbital separations for a star 10 pc distant) with a modest coronagraphic capability. NGST photometry and spectroscopy of these sources would be combined with masses obtained from the SIM and Keck interferometers to provide a rich data set on the physical and evolutionary status of massive extrasolar planets.

Fundamental Properties of Sub-Stellar Objects. NGST can provide unique astrometric information on the field BDs it detects. It should be possible to centroid the diffraction-limited beam of \sim 0.12'' at 5 μm to better than 1%, or roughly 1 mas. This would suffice to make parallax measurements from which one could obtain the distance necessary to determine absolute physical

properties and to calibrate evolutionary tracks. Further, the small mass of BD's means that an invisible object of still smaller mass will induce large reflex motions in the motion of the BD. NGST should be able to detect an astrometric signal for objects of \sim 10 Earth masses orbiting a 10 M_J BD. Since only NGST can detect these objects (they will be undetectable by a visible light instrument like SIM or by a background-limited IR interferometer like Keck), it will be important to make follow-up observations of this kind with NGST itself.

Composition and Physical Conditions. The MIR is a very useful range for examining and characterizing BD and extrasolar giant planet (EGP) structure and composition. The continuum across this region is provided by the collision-induced absorption of molecular H₂ and possibly by the absorption/scattering of particulates in the atmospheres of these bodies. The H₂ continuum will provide the best determination of the temperature structure in the upper atmospheres/photospheres of these bodies. Direct constraints on the temperature structure will link to various equilibrium and evolutionary models of these gas giants and thus constrain the size and mass of these bodies, using comparisons with models (e.g. Fig. 26). While BDs also show the 2.2 μm H₂ collision-induced fundamental feature, it is more likely that this feature is influenced by atmospheric dust particulates in the micron and submicron size range. The presence of such dust was demonstrated in the NIR by Griffith *et al.* (1998) for the BD Gliese 229B.

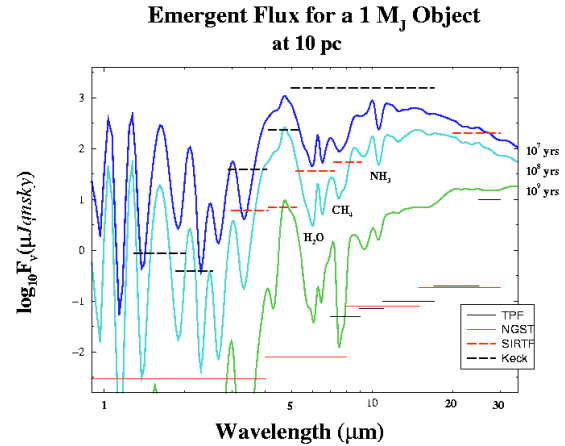


Figure 26: Emergent MIR Spectra for a 1 M_{Jupiter} object at 10 pc for ages of 10⁷, 10⁸, and 10⁹ years. (Burrows *et al.* 1997).

A variety of individual spectral features are also accessible in this spectral region. These include bands of H₂O and CH₄ in the 6–8 μm region, as well as rotational

H₂O lines longward of 20 μ m. High signal-to-noise spectra of these features in several objects will determine C/H and O/H abundances and establish the basis for models in which these constituents change with the mass and effective temperature of the object. In addition, there is a band of CH₃D whose detection would verify the expectation of deuterium depletion in BDs and differentiate between BDs and planets. Other spectral features in the 10 μ m region include NH₃. Detecting NH₃ and characterizing its abundance would either verify the temperature-dependent equilibrium expected between NH₃ and N₂ in an H₂-dominated atmosphere, or demonstrate disequilibrium chemistry. Other disequilibrium species detectable in this spectral region at $R \approx 1000$ include C₂H₂, C₂H₄ and C₂H₆. Disequilibrium chemistry has already been demonstrated by the unexpected 5 μ m detection of abundant CO in the relatively cold atmosphere of Gliese 229B (Noll *et al.* 1997). Further spectroscopy of other faint objects would characterize the extent to which disequilibrium chemistry prevails in these bodies, and the extent to which it requires the upwelling or upward diffusion of gas from the interiors of these objects.

While parts of this spectral range are accessible to ground-based observatories such as Keck, BDs are very faint and not likely to be characterized well spectrally (for $\lambda > 17 \mu$ m, even objects as bright as the now familiar Gliese 229B are too faint for detection). For BDs in multiple systems and for EGPs, major advantages of NGST in this spectral region will be its ability to resolve a secondary body from its primary down to a limit on the order of 1-2'' in the MIR, and the much reduced contrast ratio (10^{-3} to 10^{-4}) for EGPs.

3 The Instrument

3.1 Design Philosophy

The scientific goals outlined in the previous section can best be met with a MIR instrument with imaging, spectroscopic and coronagraphic capabilities, with priorities in the order given. In this section we address optical designs which provide the desired MIR capabilities, and present a viable and flexible solution. In keeping with the “proof of concept” nature of a preliminary design study, we first broke the instrument down into six conceptual “modules” which meet all of the scientific goals, and optimized the designs of these modules individually. With budgetary issues in mind, we also studied a second instrumental configuration corresponding roughly to a minimum descope option. There are a range of tradeoffs between these two limiting cases, which lead to a multi-dimensional param-

eter space in which solutions can be sought. The various options, several of which can be described naturally as upgrades and descopes, will be touched upon as needed. Several options concern detector passbands, and these are also addressed.

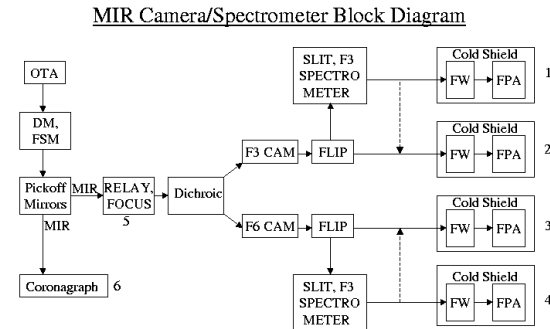


Figure 27: Block diagram of our MIR instrument concept, showing the six component modules. Note that the dichroic allows for the simultaneous use of both cameras or both spectrometers to view the same field of view. The baseline plan calls for separate FPAs with each of the five principle modules, but the dashed arrows indicate a possible tradeoff: two spectrometer FPAs for an extra pair of flip mirrors and somewhat more complicated fold optics.

To meet the triple goals of imaging, spectroscopy and coronagraphy, the most obvious starting point is three instrument modules, one for each of the functions. However, because the full wavelength coverage is approximately 5 to 30 μ m, or 2.5 octaves, optimal sampling of the focal plane calls for at least two plate scales. Thus, our baseline design calls for two camera modules and two spectrometer modules (hereafter long- and short-wavelength modules). In addition, a single coronagraphic module has been included as a separate module in the conceptual phase (although it can be subsumed into the camera modules in practice - see below). Because the desired field of view for coronagraphy is small, there is no need to split the passband in this case. Finally, to quickly depart the crowded telescope focal plane area, and to provide the option of a focus mechanism, we include an initial 1:1 imaging relay as our sixth module. A functional block diagram of our instrument design is given in Fig. 27. Note that a dual band MIR camera design provides the flexibility to match to the various detector options. As our baseline plan, we assume Si:As detectors

for all modules. However, Si:Ga or HgCdTe detectors are also possibilities for the short wavelength module. Our “minimum descope” design consists of just a single camera/spectrometer pair with an integrated coronagraph, and a single focal plane array (FPA).

3.2 The Camera Modules

At first sight, the scaling of the diffraction point spread function with wavelength might suggest that a common image quality specification is easier to meet at the longer MIR wavelengths than in the NIR. However, as a result of the likely wavelength-independence of focal plane array (FPA) pixel sizes, resulting from the use of a common multiplexer architecture (we assume $18\ \mu\text{m}$ pixel spacing here), longer wavelengths will require smaller focal ratios to match the PSF to the pixels. For example, Nyquist sampling in the $5\text{--}10\ \mu\text{m}$ range calls for an F/# of about 6–7, while in the $15\text{--}30\ \mu\text{m}$ range, an F/# of about 3 would be optimal. Thus the first challenge which a wide-field MIR camera design must face is the maintenance of high Strehl ratios across the field at small F/#’s (and in an off-axis design in the reflective case).

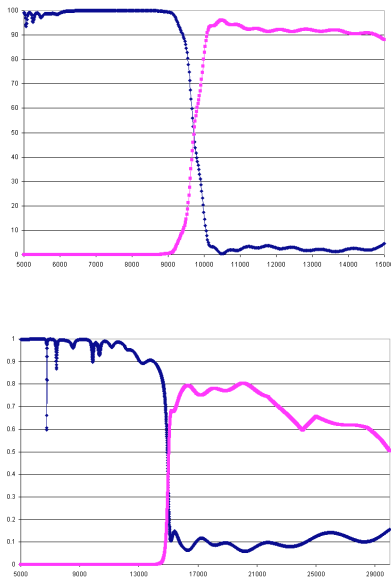


Figure 28: Dichroic beamsplitters designed for splitting the MIR passband. The beamsplitter transitions correspond roughly to the long wavelength cutoffs of HgCdTe and Si:Ga detectors.

In our baseline system of Fig. 27, we consider an F/3 system for the long wavelengths ($15\text{--}30\ \mu\text{m}$), and F/6 for the $5\text{--}15\ \mu\text{m}$ band. The plate scales per pixel, and the fields of view on a single 1024×1024 MIR array are then

$0.07''$ and $70''$ for the F/6 case, and $0.14''$ and $140''$ for F/3. For maximum efficiency, the two cameras should be able to simultaneously image the same field. To accomplish this, two things are necessary: an efficient dichroic, and a 2×2 array of FPAs in the short wavelength camera (as well as dual sets of readout electronics).

To enable the first aspect, we have explored dichroic beamsplitter designs. As can be seen in Fig. 28, two passbands can be observed simultaneously with high efficiency, but the efficiency degrades at the longest wavelengths. However, commercial coaters have indicated an expectation of improvement in beamsplitter design and performance at the longer MIR wavelengths in the next few years. This is an area in which further development is thus warranted.

The second issue concerns the differing fields of view in the two camera modules resulting from the different F/#’s. To provide matching $140''$ fields in both camera modules (and the maximum scientific output), the number of detector arrays in the short wavelength module can be increased to a mosaic of 2×2 1k arrays. Indeed, independent of these plate-scale and passband issues, a 2×2 mosaic of detector arrays covering at least the shorter MIR wavelengths would be a very valuable extension of the basic strawman concept. As the cryocooler will likely have excess unused capacity (up to about 8 MIR arrays can be cooled with an 11 mW cooler; Appendix), a wider field is certainly the simplest upgrade path from the basic single-array MIR strawman concept. In the most ambitious layout (Fig. 27), as many as 8 FPAs would be needed.

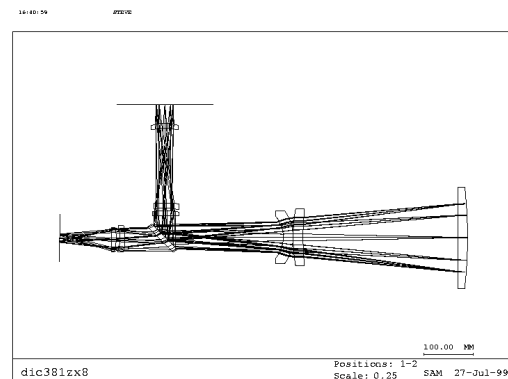


Figure 29: Refractive design for a dual MIR camera. The first beamsplitter shown in Fig. 28 is used to split the 5 to $15\ \mu\text{m}$ band.

The second challenge for MIR camera designs lies in achieving full wavelength coverage with available mate-

rials. The next question to address is then the superiority of reflective vs. refractive camera designs. Beyond about $15\ \mu\text{m}$, non-hygroscopic transmissive dielectrics become quite rare, so longward of $15\ \mu\text{m}$ an all-reflective design is a must. However, for the shorter wavelengths, we briefly examined the refractive case, and found that with a somewhat more complicated design (Fig. 29) than the straw-man case, a refractive approach could indeed achieve high Strehl ratios ($> 85\%$) across the $5\text{--}15\ \mu\text{m}$ range.

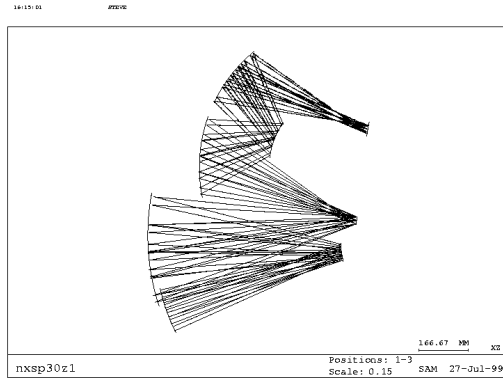


Figure 30: Overview of the optical concept: a 3-element camera (bottom; first surface at lower left), followed by a 3-element spectrometer (detector array at top right).

However, in addition to the more limited passband, two other issues also weigh in against refractive designs. In refractive cameras, the typically high dielectric indices of nonhygroscopic MIR materials makes good broadband AR coatings a necessity. AR coating designs were briefly examined, and it was found possible to achieve the necessary performance levels on individual surfaces, but as the number of lens surfaces in our acceptable solutions is large (10), cumulative losses and ghost reflections are difficult to avoid. However, more importantly, the lack of athermalization in designs based on multiple dielectrics, combined with the complete achromaticity of all-reflective designs, leads to a strong preference for an all-reflective design for a space-borne instrument. Diamond turning of all of the optical elements then becomes an ideal manufacturing approach at these long wavelengths, as in the reflective case the optics, mounts and optical bench can all be manufactured out of a common Aluminum material. The all-reflective case thus provides for a very straightforward realization, which can be implemented with current technology.

We next pursued a variety of all-reflective camera designs, for both the F/3 and F/6 configurations. The basic

solution arrived at from two somewhat different directions is the three element relay (Fig. 30). In one approach, a satisfactory three element conic-section (Korsch) design was found (Strehl ratio, S , $> 84\%$), while in a second approach, conics were abandoned, and somewhat higher Strehls were attained ($S > 92\%$). An initial 1:1 Offner relay was assumed in the second case (Fig. 27), in order to quickly move the beam away from the crowded telescope focus, and to provide a focusing capability, but otherwise this initial relay performs no vital optical function (and so was not included in the other design). It may be possible to remove the relay once the crowded telescope focal plane is better defined. The optical layout for our dual-module, all-reflective camera configuration is presented in Figs. 31 and 32, with the rays in one camera turned “on” in each figure.

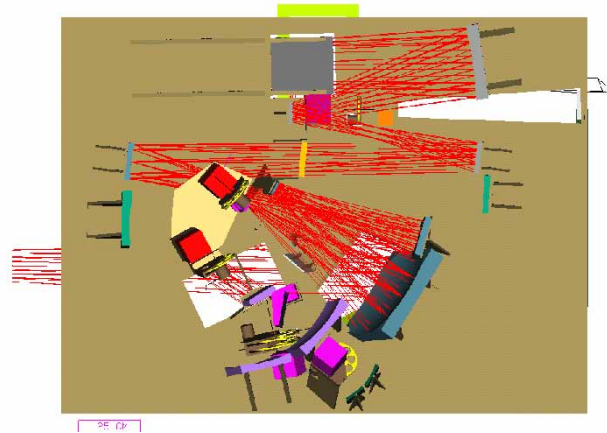


Figure 31: The F/3 long-wave MIR camera. The rays from the telescope arrive from the left, below the optical bench pictured, and pass through the bench under the grey fold mirror near the top. The 3-element Offner relay can be seen at the top of the figure, the yellow beam-splitter in the middle, a field element at the left, and the 3-element camera near the middle of the bench. Finally, a filter wheel and FPA package (red rectangular box) are seen. The FPA package and filter wheel are mounted on a thermally isolated 18 K table.

A flat pickoff mirror in the telescope focal plane is located below the plane of Figs. 31 & 32, and directs the incoming light through the imaginary optical bench pictured (Fig. 33). (The optical bench shown is an artificial element employed simply to block the ISIM mirrors from view; it is not intended as a flight structural element). The plane of this imaginary bench is parallel to the principle optical plane of the camera, and also (Fig. 34) to one of the ISIM side panels (one of the sides orthogonal to the NIR

camera mounting plane). After the beam passes through the bench, the three element Offner relay provides a relayed focal plane at unity magnification.

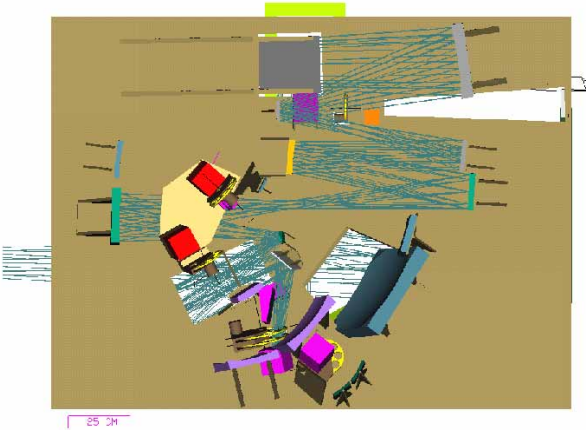


Figure 32: The F/6 camera. The rays again traverse the Offner relay, but reflect off the beamsplitter to a field element, and thence into the lower camera. The FPA package is represented by the purple rectangular box.

After the Offner (Figs. 31 & 32), the dichroic beam-splitter splits the light in the F/24 beam. (The dichroic beamsplitter could be removed if adjacent fields are used and the input optics are duplicated, but this is an inefficient approach both optically and observationally). The cameras themselves consist of field elements followed by three-element relays, consisting entirely of aspheric elements in this design (and of conics in our alternate design). Filter wheels (for broadband filters and grisms) lie in a pupil just prior to the FPAs. The filter wheels are the only unavoidable moving mechanisms in the camera modules. The filter wheels are also the location at which the volume enclosed by the innermost cold shield (at 18 K; see below) is entered. As observation of redshifted lines is not compatible with fixed filters, a selection of roughly 10 broadband filters ($R = 3$ to 10) is envisioned per camera module, targeted primarily at the broadband continuum, and at broad dust and brown dwarf spectral features.

Lightweighting of both the framework and the mirrors is planned, and Fig. 35 shows a structural design which weighs in at only 60 kg. Including the structure, optics, mounts, etc., the estimated total weight for the entire six module MIR camera/spectrometer package is 125 kg (including the cold ends of the coolers). The linear size of the framework is $1.9 \text{ m} \times 1.6 \text{ m} \times 0.7 \text{ m}$, although the last dimension is more typically under 0.5 m across the bulk of the instrument (Fig. 33).

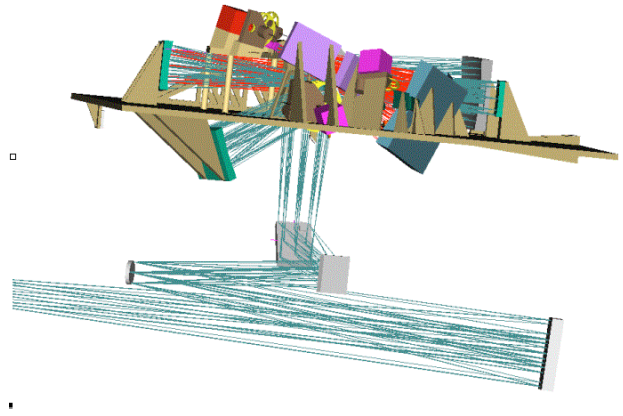


Figure 33: Side view of the camera/spectrometer assembly, showing the ISIM mirrors underneath.

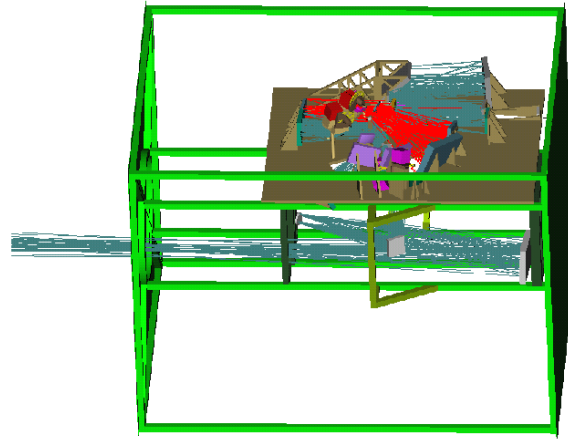


Figure 34: The MIR instrument in the ISIM.

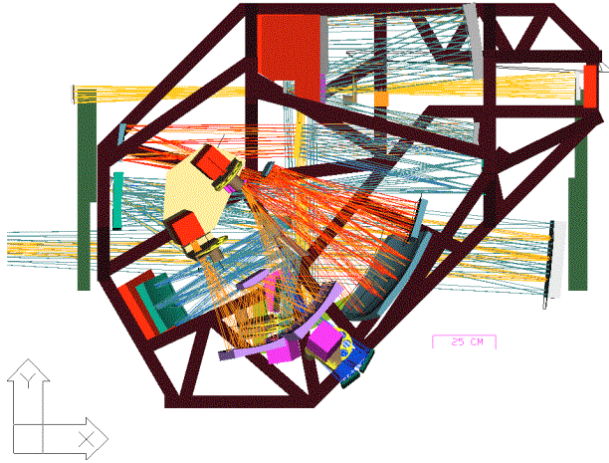


Figure 35: The MIR instrument mounted on a lightweight framework structure. The orientation is the same as for Figs. 31 & 32, but the incoming beam and the ISIM mirrors are now also seen through the structure.

3.3 Grism Spectrometers

Many options exist for broadband MIR spectroscopy, including grism, grating, and Fourier-transform spectrometers. While a Fourier transform spectrometer is best matched to the goal of equal signal-to-noise ratio (SNR) spectroscopy over a 2-dimensional field, maximum point-source sensitivity in the background-limited case calls instead for a 1-dimensional dispersive spectrometer (wherein one of the array's two dimensions is used to provide continuous measurement of the full spectrum). Thus, for reasons of higher inherent sensitivity, this study focuses solely on dispersive spectrometers.

One simple spectroscopic capability can be included directly within the camera modules themselves: grisms on the camera's filter wheel can provide a low resolving power ($R = \lambda/\Delta\lambda \approx 100$) spectrometer. For dark currents at the $1 \text{ e}^- \text{ s}^{-1}$ level (Stockman 1997), such a resolving power would lie roughly at the cross-over between the background-limited and dark-current-limited cases, for a wavelength of $10 \text{ } \mu\text{m}$ (§3.7), and so such resolving powers may be about optimal from a SNR perspective at this particular wavelength. However, the transition between the dark-current dominated and background dominated regimes is a function of wavelength (§3.7; Fig. 44), and it may be possible to achieve lower dark current performance, making the resolving power at which this transition occurs a variable rather than a constant.

Our optical analysis of the grism case reveals that for grism spectra covering an octave of bandwidth, a high

Strehl ratio can be provided across the full 2-d field up to resolving powers of order 50 to 100. Such a resolution would provide a perfect match to the first-order identification of the deepest galaxy spectra based on very broad features such as PAHs and redshifted CO bandheads. Candidate grism materials for at least the shorter MIR wavelengths include ZnSe, Ge, and CdTe, but little work has been done in actually implementing such devices. At the longer MIR wavelengths, the available materials decrease in number, but at least one non-hygroscopic dielectric (CdTe) remains available.

Two grism modes which can be envisioned: a slitless 2-d spectrometer, and a 1-d slit spectrometer. A slitless MIR spectrometer can only be considered in the low-background space environment, because of the overlapping background spectra. It would also be the simplest possible spectrometer, requiring no modifications to the basic camera beyond a slight increase in the number of filter wheel slots. There is thus little reason not to include a slitless grism capability in the camera modules.

However, this is not the most sensitive case. In the slitless case, if the noise is background-dominated, the background level is determined by the background in the full filter passband, and not in the grism resolution element. In this mode the sensitivity to point sources is thus worse than expected from the resolving power alone, by the square root of the bandwidth ratio, r (this is of course true for white noise; here we ignore spectral dependencies). However, as this mode provides spectroscopic capabilities across a (non-crowded) 2-dimensional field, the comparison is not entirely fair: in the background-limited case, the sensitivity for slitless 2-d spectroscopy of a given field is higher than that of a 1-d slit scan of the same field by the factor $(n/r)^{1/2}$, where n is the number of slit positions needed. Combined with the absence of slit losses and pointing overheads in the 2-d case, the sensitivity over a 2-d field can be roughly an order of magnitude higher in the slitless case. Thus in cases where spectroscopy over a (non-crowded) 2-d field is desired, such as e.g., “deep” galaxy fields, galaxy clusters, and the fields immediately surrounding nearby stars, the slitless mode is an attractive option. This is especially true at the shortest MIR wavelengths, as the foreseen FPAs may very well be dark-current limited for even moderate resolutions (§3.7). In the dark-current limited regime, the sensitivity advantage of a slit is gone. Nevertheless, for the ultimate in background-limited sensitivity on single objects (at the longer wavelengths), and of course in crowded fields, a long-slit mode remains essential, and so a 1-d slit grism mode is included in our baseline camera. Slit insertion is a slight complication of the camera design, as it introduces

a cryogenic mechanism, but it is essential for achieving the ultimate performance in what is likely to become a workhorse mode.

3.4 Grating spectrometers

Many of the MIR science goals can only be achieved through spectroscopy at higher resolution ($R = 1000 - 3000$) than a grism can supply. Our design therefore includes two grating spectrometer modules, corresponding to the two camera modules. Designs based on three-element, unity-magnification Offner relays, with convex gratings as the middle element, were found able to meet the requirements on image blur. Our designs place the spectrometers at the outputs of the two cameras (Figs. 30, 36 & 37), with flip mirrors located just prior to the final pupil image to redirect the light into the spectrometers. The existing cameras thus take care of the plate scale reduction. JPL has considerable experience with such convex gratings, and they are obtainable commercially.

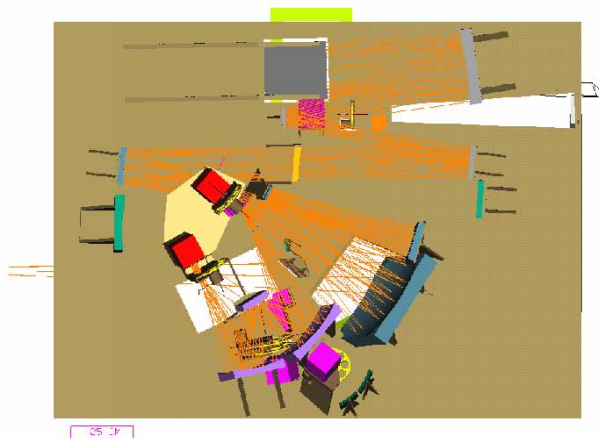


Figure 36: The F/3 spectrometer. First the Offner relay and the F/3 camera are traversed to reduce the F/#. A flip mirror just prior to the pupil at the camera's filter wheel redirects the light toward the 3-element spectrometer.

Many spectrometer configurations are possible, the simplest being a long slit spectrograph. Integral field spectrometers are also of considerable interest in the MIR. With Nyquist sampling of a spectrum along the full length of a $1k \times 1k$ array, and an instantaneous spectral coverage of roughly an octave (to match to typical grating efficiency curves), resolving powers of order 500 can be achieved without the need for a fine grating drive mechanism. Positioning two parallel, offset slits at diametrically opposite sides of the array could provide for a doubling of the resolving power in straightforward fashion. As the

telescope can be used to position the star sequentially on both of the two slits, no slit-changing mechanism would be needed in the simplest case. Thus, because of the large FPA format, a resolving power of order 1000 can be obtained in a fairly simple instrument without either a fine grating drive or a slit mechanism.

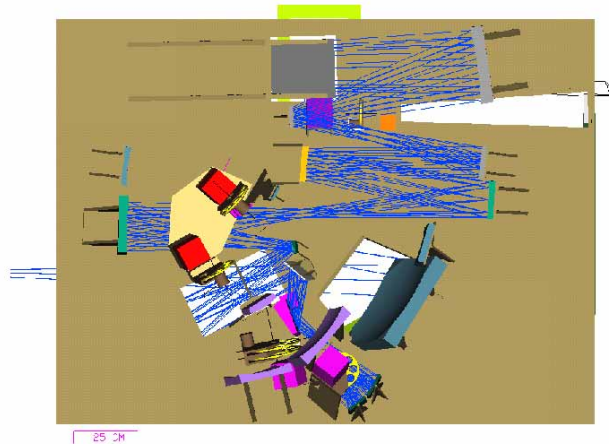


Figure 37: The F/6 spectrometer. First the Offner and the F6 camera are traversed to reduce the F/#, and then a flip mirror prior to the filter wheel redirects the light to the spectrometer.

As mentioned, this configuration relies on flip mirrors to select between the camera and spectrometer modules, and so these cryogenic mechanisms are needed in this design (unless the spectrometer fields were offset from the camera fields, but this would call for duplication of the reducing optics). A variety of slit widths, and so a slit-change mechanism, might also be considered, but given the large FPAs available, the need to change slits can likely be avoided by design. For example, given the large number of pixels, a single long slit would be overly long. However, with a tapered slit, a source could be positioned on the slit as desired, eliminating the need for multiple slits. In addition, an image slicer covering only a part of the focal plane could be accommodated in concert with tapered slits. In this scenario, part of the array would provide an integral field capability, while another would be dedicated to long slits, with a minimum of moving parts.

For somewhat higher resolving powers, either a grating drive, a cross-disperser (about three orders per octave), or the sequential measurement of several orders (in concert with order sorting filters and an appropriately blazed grating) are good possibilities. (Immersion gratings are still fairly limited in terms of proven performance, so these are not considered here). In order to avoid a cryogenic, space-

qualified grating scan mechanism, we examined both of the latter two possibilities, and found both to be fairly straightforward to implement optically. The sequential measurement of orders provides for the simplest design, but is less efficient than the cross-dispersed case.

However, while a cross-disperser is desirable for obtaining the highest spectral resolution, it is important to remember that at the higher resolving powers, the MIR FPAs may be dark current limited over a significant fraction of the band (the short end). Use of the minimum acceptable resolving power will thus be vital for maximizing SNR, and in particular, there is no reason to push beyond $R \approx 1000$ in the dark-current limited case ($R > 3000$ is not called for by the science program in any case). Thus, at the long wavelengths, where one is likely to be background-limited, optimization of the science return and SNR occurs with a (reflective) cross-disperser configuration. This will only be the case at the shorter wavelengths if the FPA dark currents are reduced by about a factor of 50 from the strawman level of $1 \text{ e}^- \text{ s}^{-1}$. Our baseline spectrometers therefore consist of $R = 3000$ cross-dispersed systems for both the long and short wavelength modules, in the expectation of reduced dark current performance (§3.7). However, to optimize the short wavelength SNR, this spectrometer module also includes a simpler direct $R = 1000$ spectrometer (the cross-disperser is replaced by a mirror, and a different grating is selected on a turntable). These modes then call for inclusion of grating selectors (with large rotations to stops), as well as a grating/mirror flip mechanism for the short-wavelength cross-disperser/mirror selection. Obvious descope options/simplifications are first to abandon cross-dispersion by going to sequential measurement of multiple orders (with prefilter order-sorting), and then as a next step, to retain only the direct $R=1000$ mode. Again, this latter mode can be implemented quite simply.

3.5 Coronagraphy

Because of the importance of MIR coronagraphy to near-neighbor detection and characterization (§2.7), our instrument includes a coronagraphic mode. There are several possible implementations. First, to minimize scattered light, a minimum number of reflections is desirable. As the field needed is not very large, we have thus included a very basic F/24 coronagraph in our design (Fig. 38), with the simplest possible optical layout (no plate scale reduction, no relay). The optical train includes only two powered elements (the Offner relay is bypassed with a separate pickoff mirror), and provides for the placement of a fixed focal plane mask and a 2-axis adjustable Lyot stop.

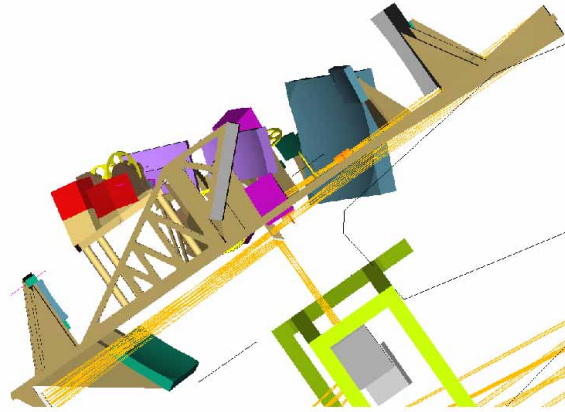


Figure 38: The coronagraph (in yellow), consisting of a fixed focal plane blocker, a spherical mirror, a (2-axis) translatable Lyot stop, a second powered optic, a filter wheel, and an FPA.

However, as such a coronagraph would introduce another complete camera module and FPA into the system, a second option is to implement the coronagraph directly in the existing cameras. The already existing relayed focal and pupil planes can then be used for insertion of the pupil and focal-plane masks, at the cost of some degradation in performance. Indeed, it is mainly the short-wavelengths ($5 - 10 \mu\text{m}$) which call for coronagraphy (due to the smaller beams and higher sensitivities), and therefore only the short wavelength MIR camera would need a coronagraphic capability. Furthermore, by going through the camera modules, the light could then be sent onward for spectrographic analysis, providing for spectroscopy of near neighbors. The combined camera/coronagraph option for the short-wavelength MIR camera module thus has significant benefits, and so this is the approach taken in our minimum descope option. Finally, if a facility NIR coronagraph becomes part of NGST's instrument suite, inclusion of a dichroic beamsplitter and an MIR FPA in such a package would also be a viable option.

3.6 Detectors

Table 2 lists the detector materials with the greatest promise for meeting the demanding performance levels required by NGST. The doped silicon materials (Si:X) all assume a blocked impurity band conductor (BIB) structure, while the mercury cadmium telluride (HgCdTe) is a more traditional photovoltaic structure. The quantum well interference photodiodes (QWIPs) are a relatively recent

development. They have inherently narrower passbands (10–50% is typical), lower quantum efficiencies (20% maximum), and higher dark currents, but they may provide a viable alternative if sub-30 K cooling is not available on NGST.

Table 2: MIR Detector Materials

Material	Wavelengths	Temp.
Si:Ga	4–17	14
Si:As	5–27	10
Si:P	8–35	8
Si:Sb	12–45	5
Hg _{1-x} Cd _x Te (x=0.21)	4–12	30–35
QWIP(GaAs/AlGaAs)	Tunable	25–35

Of these choices, arsenic-doped silicon (Si:As) is by far the most mature. Arrays of such detectors have flown on ISO, Spirit III/MSX, and are featured prominently in SIRTf. Arrays with performance nearly adequate for NGST are already available; the primary deficiency is the array format. However, 1024^2 arrays are under development by several vendors. Operationally, the largest drawback is the fact that these arrays need to be cooled to 10 K or lower to meet the performance specifications.

Because the response of low doping, low background Si:As arrays drops precipitously by about $27\text{ }\mu\text{m}$ (Fig. 39), they are not suitable for observations of the $28.2\text{ }\mu\text{m}$ H₂ rotational ground state line. Should such observations be desirable, or if simply a longer wavelength cutoff is desired, phosphorus-doped silicon (Si:P) is an attractive option (also shown in Fig. 36). It has a response which extends to $35\text{ }\mu\text{m}$ and an operating temperature of 8 K, only a 2 K penalty with respect to Si:As. However, it does have a decreased response in the critical 5–10 μm range. Antimony-doped silicon (Si:Sb) has an even longer wavelength response ($\sim 45\text{ }\mu\text{m}$), but must be cooled to 5 K or lower to meet the dark current requirements. Si:P is also significantly easier to make than Si:Sb as it has good gas phase carriers for deposition during epitaxy. Gallium-doped silicon (Si:Ga) is attractive if the coolers cannot provide sub-10 K cooling but can provide 10–14 K levels. The cutoff wavelength for Si:Ga is approximately $17\text{ }\mu\text{m}$, which is adequate for a good fraction of the planned MIR extragalactic science, but not for the majority of the galactic science.

HgCdTe or QWIP detectors will be required if cooling below 30 K is deemed infeasible. Each can be made to cover at least the 5 to $10\text{ }\mu\text{m}$ range with acceptable performance, thus meeting the desire for NGST to extend to “10 plus” μm . Here again, a very large fraction of the science

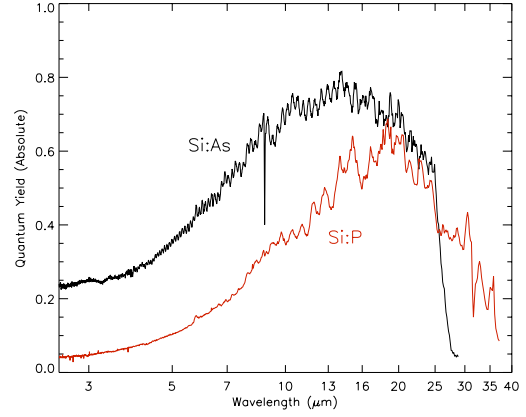


Figure 39: Spectral responses of Si:As and Si:P detectors compared. The Si:As curve gives the absolute response of a low-doping “low-background” Boeing array. The Si:P curve is for a non-optimized detector.

topics discussed earlier would be lost.

The detector performance required to meet the scientific goals outlined in this report can be summarized as follows: size $\geq 1024^2$; wavelength range 5–30 μm , quantum “efficiency” $\geq 70\%$; dark current $< 0.08\text{ e}^-/\text{s}$ for $R = 100$ spectroscopy and $0.02\text{ e}^-/\text{s}$ for higher resolution spectroscopy; read noise $< 5\text{ e}^-$. The size and spectral range requirements follow directly from the science discussion; the other are discussed below.

3.7 Sensitivity

To compute the background fluxes and instrument sensitivities one might expect with NGST, a model was developed which assumed the following properties: an 8-m filled aperture telescope, f/6.5 reimaging optics, $18\text{ }\mu\text{m}$ detector pixels, or a plate scale of $0.07''/\text{pixel}$, $0.4\text{ e}^-/\text{s}$ dark current, and read noise \ll detector current noises. (This dark current value is derived from current performance levels of 128×128 pixel FPAs, scaled to $18\text{ }\mu\text{m}$ pixels, at an operating temperature of 10 K. These values can be further improved both by materials development in the next several years, and by lowering the operating temperature by a degree or two). Mirrors are assumed to have 98% reflectivity. The ISIM is at 40 K and has 4 reflections before the camera including the final pick-off mirror. The camera is assumed to have 10 reflections. The detector enclosure temperature is a variable, and the final filter, with 85% transmission, is the only optical element at this temperature. Scattered light from sun shield is a 93 K graybody with an “emissivity” of 5×10^{-5} , and the

zodiacal background is represented by a 266 K graybody with emissivity = 10^{-7} through the plane of the Solar system, and 1/3 of that at the ecliptic poles. The spectral resolution varies as indicated, but is 10% ($R=10$) through the bulk of this discussion. Primary mirrors with a uniform temperature of 50 K and with a “gradient” from 40–60 K (approximated by zones of 40, 45, 50, etc. K), along with a 75 K primary (which is the hottest mirror consistent with “zodi-limited at 10 μm ”) were also evaluated. Most of these assumptions follow from earlier NGST studies.

The model attempts to be as rigorous as possible given the current state of information about the NGST configuration. For example, measured quantum yield curves for the detectors were used and integrated over to determine photon fluxes. On the other hand, square passbands were assumed for the filters. Broadly speaking, the model (written in the IDL data processing language) follows these general steps:

1. Compute the spectral response of the detectors and atmosphere (for comparison) over finely sampled wavelength intervals (roughly $R=5000$).
2. Compute the optical throughput at various points through the instrument
3. Compute the background photon flux per pixel per unit wavelength from various optical elements. This is given by $f(\lambda) = \epsilon A \Omega B(\lambda, T) / h\nu$ where A is the detector pixel area, Ω is the solid angle viewed by that pixel, ϵ is the emissivity of a given element, ν is the frequency, and $B(\lambda, T)$ is the Planck function.
4. Integrate the product of the background flux and the detector quantum yield over all wavelengths, to yield electrons per second per pixel from each optical element. For all elements closer to the array than the filters, the form is $\int_0^\infty f(\lambda) q(\lambda) d\lambda$ where $q(\lambda)$ is the quantum yield of the detector. For elements more distant than the filters, the equation becomes $\int_{\lambda_1}^{\lambda_2} f(\lambda) q(\lambda) d\lambda$ for each filter of interest, where λ_1 is the short wavelength cuton, and λ_2 is the long wavelength cutoff (defined as the central wavelength plus/minus half the spectral resolution, *e.g.*, for a 10.0 μm filter and $R=10$, the values are 9.5 and 10.5, respectively).
5. Sum the contributions from all elements to determine the total background photocurrent generated in each pixel. Plot this as a function of various filter central wavelengths. Note that because of the smoothing implicit in the calculations, sensitivity at a given

wavelength may be implied even if the detector quantum yield curve shows none, due to the broadening effects of the filters. For example, a 10% filter centered at 28.2 μm would cover the range 26.8 to 29.6 μm , and so sensitivity arises from the 26.8 to 28.0 μm range even though there is zero quantum efficiency at the passband center.

6. Compute the detected photon flux from a star of magnitude 0 by integrating the throughput over each filter passband and the respective detector quantum yield.
7. Compute the approximate number of pixels contained within the PSF, assumed here to be a square $2.44\lambda/D$ pixels on a side.
8. Find the stellar photocurrent required to yield a signal-to-noise of 1 in 1 second when summed over that number of pixels. Invert back to Janskys, then scale to 10σ in 10,000 seconds.

The results of our series of calculations are presented in the final five figures. Fig. 40 shows the background contributions for the various NGST elements and the MIR camera integrated over an $R=10$ passband. Ignoring initially the detector enclosure curves, dark current (at the assumed 0.4 e^-/s level) dominates at wavelengths $< 5 \mu\text{m}$, and zodiacal light dominates for 5–10 μm . A hot primary (75 K) would dominate all other sources at wavelengths $> 10 \mu\text{m}$; with a 50 K (either single temperature or gradient) primary, the scatter from the 93 K sun shield off the primary and secondary into the beam could dominate if the surfaces were rough enough. In either case, the telescope would likely limit sensitivities at $\lambda > 18 \mu\text{m}$ no matter the configuration, unless it could be entirely passively cooled to < 40 K. ISIM and camera optics contributions are small compared to that of the telescope. Our results to this point are thus consistent with earlier NGST studies.

Also shown in Fig. 40 is the contribution of the detector enclosure, for various assumed temperatures. At 18–20 K, the effects of the enclosure begin to be seen at the shortest wavelengths. At $\lambda < 10 \mu\text{m}$, the enclosure completely dominates all other background sources for enclosure temperatures above 22–23 K. The reason for this is that the photon flux from such an enclosure is a constant, independent of whatever light is impinging on the detectors along the optical axis, and so is analogous to a constant level dark current. This contribution may be computed by integrating the product of the Planck function and the unfiltered detector response (roughly 2–27 μm for low doping Si:As), over 2π steradians. The colder the enclosure (or the shorter the wavelength cutoff of the detector), the lower this background. Overall sensitivity is

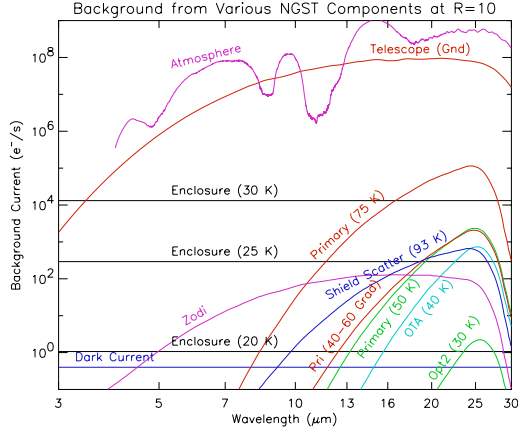


Figure 40: Background contributors for an R=10 passband. For comparison, atmospheric and telescope emission are given for an 8 m groundbased telescope. With a cold detector enclosure, dark current dominates for $\lambda < 5 \mu\text{m}$, zodiacal light for $5\text{--}10 \mu\text{m}$, or $5\text{--}18 \mu\text{m}$, depending on configuration, and the OTA for $\lambda > 18 \mu\text{m}$. As the detector enclosure temperature is raised from 18 to 25 K, the enclosure emission gradually comes to dominate all other background sources out to $\lambda \approx 12$ to $18 \mu\text{m}$.

affected in a manner exactly analogous to dark current; that is, the quietest regions, *i.e.*, the *shortest* wavelengths are impacted first, and most severely.

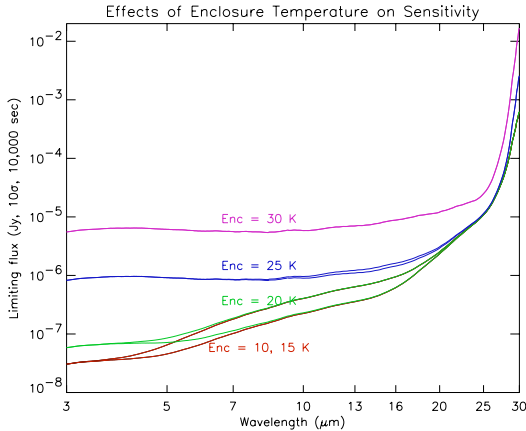


Figure 41: The sensitivity of NGST with a 50 K primary mirror (no sun shield scatter), but where the enclosure temperature is varied. For enclosure temperatures $\geq 20\text{--}25 \text{ K}$, the zodi limit is not reached. The steep rise at $30 \mu\text{m}$ is due to the Si:As response falloff.

Fig. 41 shows the sensitivity degradation for R=10 as

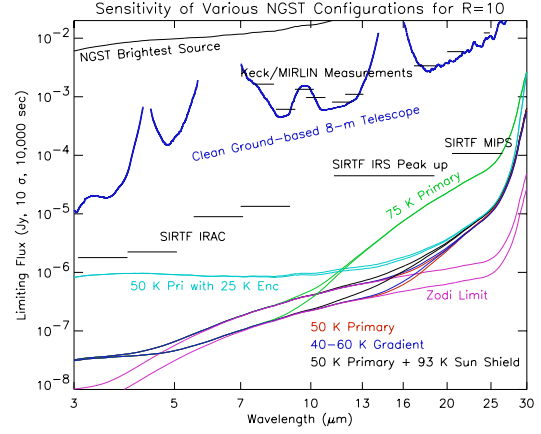


Figure 42: Sensitivity contributors for the various telescope configurations of Fig. 40, at R=10. NGST will be over three orders of magnitude more sensitive at $10 \mu\text{m}$ than a highly optimistic estimate for a ground-based 8-m telescope. The “brightest NGST source” was computed by finding the flux from a point source which would fill the well at the central pixel of the PSF to 85% in the minimum 1 sec readout time.

the enclosure temperature is varied from 10 to 30 K (for an otherwise dominant 50 K primary). It is clearly desirable to keep the enclosure temperature as low as possible—at the detector temperature is of course desirable, but a temperature of up to about 18 K will not impact performance at the shortest MIR wavelengths.

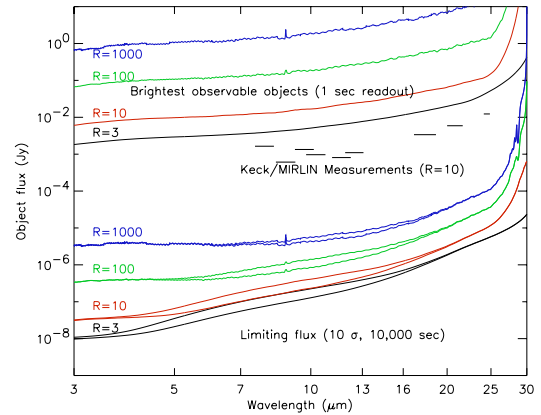


Figure 43: Sensitivity and saturation flux curves for various resolving powers.

Fig. 42 shows our complete sensitivity model in terms of incident flux. The bottom two curves in violet are the zodi-limit (*i.e.*, telescope, all optics, and the innermost en-

closure are at absolute zero). The upper violet curve is through the plane (“emissivity” of 10^{-7}), the lower along the pole (3×10^{-8}). The 50 K family of curves are all roughly the same. The model with the 93 K sun shield makes things a bit worse, particularly along the pole, but not dramatically so. The 40-60 K gradient mirror is not significantly different from an isothermal 50 K mirror. The short wavelength deviations from the zodi limit are due to the detector dark current and a warmer than optimal detector enclosure. The degradation that a 25 K detector chamber would engender is quite evident: one to one and a half orders of magnitude in sensitivity are lost in the 5–10 μm range with the warmer enclosure.

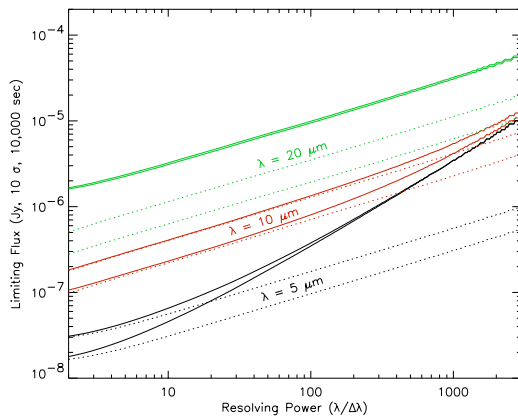


Figure 44: Instrumental sensitivity as a function of resolving power for three wavelengths. For each wavelength, the minimum and maximum zodi-limited performance are represented by the dotted lines; the predicted system performance is shown by the solid lines. At 10 μm the system is zodi-limited only for $R < 100$. Sensitivities at higher resolutions are degraded by dark current. At 5 μm , the system is essentially always dark current limited. At 20 μm , the sensitivity is always telescope-limited.

Finally, Fig. 43 shows the expected sensitivity and saturation flux (see caption of Fig. 42) for various spectral resolving powers, under the “cold detector enclosure” assumption. At the short wavelengths, dark current becomes increasingly important as the resolving power is raised. This can be seen even more clearly in Fig. 44, which plots the sensitivities as a function of resolving power for three wavelengths. At 20 μm , the sensitivities are always background limited to $R = 3000$, but by the telescope rather than zodiacal background. At 10 μm , the system is zodi-limited at resolving powers of up to about 100, and dark current limited thereafter. However, at 5 μm , the system is dark-current limited down to resolving powers of order

3. Thus, optimal performance in the critical 5 to 10 μm range requires dark currents more than an order of magnitude lower than the assumed baseline, and so this should be re-evaluated in light of current prognostications.

3.8 Electronics

We anticipate that the detector arrays will have 4 analog outputs per array. These will be buffered by high stability preamps located near the MIR instrument in order to drive the signals to the A/D converter bank in the ISIM common electronics area. Because the dynamic range of these arrays will be quite high (of order 20,000), 16-bit converters will be required to adequately sample the signal. The maximum internal data rate (assuming 2 cameras, or 5 arrays are read simultaneously) is then 0.1 Gbit/s. This digitized signal will be sent to a coadders/memory area, where frames (potentially thousands) can be co-added to build up signal-to-noise before downlinking.

The readout rate of low doping MIR arrays is likely to be of order 1 second, which sets the minimum integration length at 1 second, without subarraying or extended reset control. For coronagraphic observations of bright stars, faster readouts are highly desirable to avoid saturation—thus, subarraying is a necessity for this case. Faster readout of the entire array may be possible, but will increase the power dissipation of the array. The longest on-chip integrations will be set by the cosmic ray flux; a single integration should end before the number of cosmic ray hits is significant. This time is expected to be in the range 10 to 60 seconds, necessitating the need for 32-bit deep coadders if 10,000 integrations are required before downlinking. The coadders should have built-in cosmic ray detection algorithms so that an obvious hit could be removed from the data stack before it is sent down.

Assuming space qualified equivalents to the best currently available commercial technology are found, we estimate that the power dissipation for the analog preamp portion of the readout is of order 100 mW per output, while the A/D portion is likely to be about 400 mW per output. Our most complex system (8 arrays) would thus consume about 16 W for the readout electronics associated solely with our instrument. Some power is required to drive the filter wheels, but this will be a transient load of order a few Watts.

We assume coaddition, storage, and transmission will be handled by a common ISIM electronics bay; the power necessary to process our data may be near 100 W (assuming a Pentium-class processor plus memory and storage). We do not expect the total system load to exceed 150 W (excluding the coolers).

3.9 Descope Options

With our modularized approach, the descope options are clear: removal of modules (e.g. removal of spectroscopy, or reduction to 1 camera/spectrometer pair), reductions in the number of FPAs, selecting between flip mirrors and extra arrays, and shifting of detector cutoff wavelengths (detector types).

Of course, due to the radical sensitivity gain provided by NGST in the MIR, as well as the huge angular resolution gain compared to ISO and SIRTf, modest primary mirror configuration descopes would still leave the MIR science case largely intact.

3.10 Development Path

The reflective optics need little development, but the chromatic elements need a modest amount. In particular, development of MIR beamsplitters, grisms, and dielectric cross-dispersers is needed in the next several years. Convex gratings and the diamond turning of mirrors are known technology, but reduction of scatter at the shortest wavelengths is an area deserving of optimization effort.

The mechanisms share commonality (copied filter wheels and flip mirrors), and so the development effort here is to define filter wheels, and other actuators/mechanisms which will work at cryogenic temperatures. Luckily, there is a growing data base gained by other flight instruments (ISO, NICMOS, SIRTf), upon which to build.

The two most critical development needs are the detector arrays and cryocoolers. Substantial funding and development effort will be required in these two areas in the next few years. Detector development is funded and underway, but real performance and cross-comparison data will be needed. Facilities for the performance testing of arrays and of optical components such as beamsplitters and grisms is thus essential in the fairly near term. Likewise, the readout electronics for the projected arrays requires development, preferably in concert with the manufacturers.

As part of our cryocooler section, we provide a development path for sorption coolers.

3.11 Budget

Our MIR instrument design was costed in three ways. First, a simple mass/lifetime parameterization gave a \$39M–\$85M range. Second, a detailed grass roots cost estimate (next page) came in at \$65.8M. Both of these estimates are for the full six-module instrument described.

Finally, a descoped version consisting simply of one camera and one spectrometer module, with a common FPA, was ROM costed at Ball Aerospace at \$39.5M. The different estimates are thus all consistent with each other.

However, note that these cost estimates exclude the cryocooler, which is costed separately in the appendix at \$20.8M (from Phase-A to launch).

4 References

- Armus, L., Neugebauer, G., Soifer, B.T. & Matthews, K. 1995 AJ 110 2610
- Benitez, N., Broadhurst, T., Bouwens, R., Silk, J. & Rosati, P. 1999 ApJ 515 L65
- Boogert, A.C.A. 1999 Ph.D. Thesis, Groningen
- Burgasser, A.J. *et al.*, 1999 ApJ 522 L65
- Burrows, A., Marley, M., Hubbard, W.B., Lunine, J.I., Guillot, T., Saumon, D., Freedman, R., Sudarsky, D. & Sharp, C. 1997 ApJ 491 856
- Dermott, S.F., Jayarama, S., Xu, Y.L., Gustafson, B.A.S., Liou, J.C. 1994 Nature 369 719
- Dwek, E. *et al.* 1998 ApJ 508 106
- Elvis, M., Wilkes, B.J., McDowell, J.C., Green, R.F., Bechtold, J., Willner, S.P., Oey, M.S., Polonski, E., & Cutri, R. 1994 ApJS 95 1
- Gaidos, E. PASP 1998, 110, 1259
- Gardner, J.P. & Satyapal, S. 1999 BAAS, 31, no.3
- Genzel, R. *et al.*, 1998 ApJ 498 579
- Gibb, E.L., Gerakines, P.A. & Whittet, D.C.B. 1998 AAS 193 7102
- Greaves, J. *et al.*, 1998 ApJ 506 L133
- Guilloteau, S., Dutrey, A. & Simon, M. 1999 A& A 348 570
- Haiman, Z. & Loeb, A. 1998 ApJ 503 505
- Hogerheijde, M.R., Van Dishoeck, E.F., Salverda, J.M. & Blake, G.A. 1999 ApJ 513 350
- Kirkpatrick, J.D. *et al.*, 1999 ApJ 519 802
- Langer, W.D. 1999 in Protostars & Planets IV, Univ. of Arizona Press: Tucson, AZ, in press
- Laurent, O., Mirabel, I.F., Charmandaris, V., Gallais, P., Madden, S.C., Sauvage, M., Vigroux, L. & Cesarsky, C. 1999 A& A, submitted
- Moorwood, A.F.M. *et al.*, 1996 A&A 315 L109
- Lutz, D., Spoon, H.W.W., Rigopoulou, D., Moorwood, A.F.M., & Genzel, R. 1998 ApJ 505 L103
- Mirabel, I.F., Vigroux, L., Charmandaris, V. *et al.*, 1998 A& A 333 L1
- Noll, K.S., Geballe, T.R. & Marley, M.S. 1997 Ap J 489 L87
- Oh, S.P. 1999 ApJ, in press, astro-ph/9904255
- Reid, I.N. *et al.*, 1999 ApJ 521 613

Schutte, W.A., Greenberg, J.M., Van Dishoeck, E.F., Tielens, A.G.G.M., Boogert, A.C.A. & Whittet, D.C.B. 1998 Ap& SS 255 61
 Shier, L, Rieke, M. & Rieke, G.H. 1994 ApJ 433 L9
 Stockman, H.S., 1997, The Next Generation Space Telescope: Visiting a Time When Galaxies Were Young, AURA, Inc., p.13
 Sturm, E. *et al.*, 1999, in prep.
 Thi, W.F., Van Dishoeck, E.F., Blake, G.A., van Zadelhoff, G.-J. & Hogerheijde, M.R. 1999 ApJ 521 L63
 Thompson, R.I., Storrie-Lombardi, L.J., Weymann, R.J., Rieke, M.J., Schneider, G., Stobie, E., & Lytle, D. 1999 AJ 117 17
 Tielens, A.G.G.M., Waters, L.B.F.M., Molster, F.J. & Justtanont, K. 1998 Ap& SS 255 415
 Waelkens, C. *et al.*, 1996 A&A 315 L245
 Xu, C., Hacking, P.B., Fang, F., Shupe, D.L., Lonsdale, C.J., Lu, N.Y., Helou, G., Stacey, G.J., and Ashby, M.L.N. 1998 ApJ 508 576

	GRASS ROOTS COST ESTIMATE FOR NGST MIR CAMERA					
WBS	Element	Phase A	Phase B	Phase C/D	Launch	Total RY \$K
1.0	Management	320	490	1417	660	2887
2.0	Science	360	760	3475	2760	7355
3.0	Systems Eng.	190	420	1216	314	2140
4.0	SR & QA	200	210	1060	230	1700
5.0	Structure/Mechanisms	285	1100	14220	425	16030
6.0	Optics	215	555	8690	245	9705
7.0	Electronics	205	800	4874	420	6299
8.0	Thermal Eng.	200	407	1444	180	2231
9.0	Software	190	390	1767	240	2587
10.0	Detectors	170	218	4275	110	4773
11.0	Reserves	245	802	8480	558	10085
	Total RY \$K	2580	6152	50918	6142	65792

Activity Name	2001				2002				2003				2004				2005				2006				2007			
	2nd	3rd	4th		1st	2nd	3rd	4th	1st	2nd	3rd	4th	1st	2nd	3rd	4th	1st	2nd	3rd	4th	1st	2nd	3rd	4th	1st	2nd	3rd	4th
Phase A																												
Phase B																												
Phase C/D																												
Post Deliv. Support																												
Major Milestones																												
Concept Study																												
Preliminary Design																												
Detailed Design																												
Procure/test optics/filters																												
Procure/test detectors																												
Procure elec. parts/PCBs																												
Proc./fab mech piece parts																												
Assemble/test subassemblies																												
Integration Test/Qual & Calibrate.																												

Appendix A: Cryocooler Options for NGST MIR Detectors

A.1 Status of Cryocooler Technology for MIR Applications

The nominal operating temperature for the NGST Integrated Science Instrument Module (ISIM) while in orbit will be in the range of 30 K to 40 K from passive. Since passive cooling becomes extremely ineffective below 30 K, the detectors for the mid-infrared camera option will require active cooling to lower temperatures. As shown in Table 1, there is a trade-off between the IR detector wavelength and its operating temperature range. The broadest IR frequency range will require the detectors to be cooled to 4 – 8 K, but significant extension of the cut-off wavelength is possible for cooling below ~ 15 K. The cooling loads for the <10 K detector arrays is estimated to be about 1 mW per detector array, and less than 10 mW for up to seven detector arrays, including parasitic heat loads.

There are two general approaches to produce the temperatures needed by the MIR detectors. First, stored cryogenics (i.e., liquid helium for $T \sim 4$ K or solid hydrogen for ~ 6.5 K to ~ 13 K) can be used. While such storage dewars have been used on previous space flight missions of duration less than 1-2 years, many issues involving mass, volume, lifetime, safety and risk suggest that stored cryogenics are inappropriate for NGST. The second approach is to use long life closed-cycle cryocoolers¹. Table 1 summarizes the viable choices for mechanical coolers that will produce temperatures below ~ 25 K. Although multiple-staged Stirling cycle coolers can operate below 20 K², these mechanical coolers are very power inefficient - a small Stirling cycle cooler would typically operate² with a Carnot coefficient of performance of 2% at 16 K and even less at lower temperatures. This inefficiency, common to all regeneration cycle coolers, is associated with the regenerator heat capacity. Unfortunately, the heat capacity of conventional regenerator materials also falls by T^3 as the temperature decreases making it very difficult to reach very low temperatures without using novel regenerator materials². Stirling cycle coolers are also plagued by residual vibrations over a wide frequency band. The pulse-tube variant¹ of Stirling cycle coolers significantly reduce electromagnetic interference and vibration levels. A commercial laboratory 2-stage pulse-tube cryocooler has been recently announced³ that can produce 0.5 W cooling at 4.2 K for 5.5 kW of input power. A conceptual configuration of a 4 K pulse-tube cooler that may be adaptable for space applications has been reported⁴. Disadvantages⁵ of the pulse-tube design include a slightly lower Carnot efficiency than the Stirling cooler due to irreversibilities in the pulse tube expansion (which also exacerbates the regeneration problems < 20 K) and integration difficulties due to the configuration of the cooler components and the location of the cold block. Stirling and pulse-tube coolers do not scale well for cooling capacities below 200 mW. The input power requirements for cooling to these low temperatures are higher due to poor regenerator performance and real-gas effects. Hence, the Stirling and pulse-tube coolers are presently not viable options for the NGST MIR detectors.

Three distinct types of cryocoolers remain as viable options for NGST-MIR coolers: (1) Joule-Thomson (J-T) cycle coolers with sorption compressors; (2) J-T cycle coolers with mechanical compressors; and (3) miniature Turbo-Brayton mechanical coolers. The operating temperatures and development status for different variations of these coolers are compared in Table 1. While the mechanical coolers would probably use only He as the refrigerant fluid, the sorption coolers use either helium (with a charcoal pump) or hydrogen (with a metal hydride pump) over the cooling range ~ 1 -8 K or 8–25 K, respectively. The sorption coolers can also function in either a

periodic mode (i.e., intermittent cooling for short times) or continuous operation to provide a constant cold temperature.

There are presently no space flight demonstrated, long-life cryocoolers available in the temperature range desired for MIR detectors. However, several of the coolers are in various stages of development for cooling applications below 25 K.

- 1.) Creare, Inc. is currently developing a two-stage 25K/6K Turbo-Brayton demonstration cooler⁶ under the sponsorship of the GSFC for potential NGST application. This type of cooler⁷ has been demonstrated for long-life engineering tests and space flight as a multi-watt cooler at $T > 65$ K and as a 1 W cooler at 35 K in the laboratory⁸. It remains to be demonstrated whether the Turbo-Brayton coolers can operate efficiently at low power loads and temperatures below 30 K, since parasitic heat losses impact scaling to smaller

Temp (K)	Closed Cycle Continuous Coolers																Stored	Passive Cooling
	Sorption Compressors												Mechanical		Cryogen			
	Detector					Periodic		Continuous					continuous					
	Si:Sb	Si:As	Si:Ga	HgCdTe	QWIP	Liquid helium-Charcoal	Solid Hydrogen-Metal Hydride	Liquid helium-charcoal compressor	Supercritical Helium J-T-Charcoal compressor	hydrogen, liquid w/sorption compressor	hydrogen, solid w/sorption compressor	helium J-T mechanical compressor	Turbo-Brayton	Solid Hydrogen				
3						2 He stages		2 He stages										
4						1 He stage		1 He stage				SD						
5						POP		POP				(Planck)						
6									development at JPL				development (Goddard)	SD				
7																		
8											concept only							
9							SD (BETSCB)											
10																		
11																		
12																		
13																		
14																		
15																		
16										POP								
17																		
18																		
19																		
20													POP					
21																		
22																		
23																		
24																		
25																		
26																		
27																		
28																		
29																		
30																		
31																		
32																		
33																		
34																		

Table A1. The operating temperature ranges for available mid-IR detectors along with their effective wavelengths are compared to the temperatures provided by various cooling options. The status for the different coolers are: POP = Proof-of-Principle by ground test, OTS = Off-the-Shelf for ground applications, SD = Space Demonstration.

designs. Furthermore, 20 mW T-B coolers will require reduced dimensions for critical components such as the turboalternator and heat exchanger unit, which will impose significant manufacturing challenges for machining tolerances, etc. and increased risk of instability during operation. Turbo-Brayton coolers also require that the waste heat be exhausted near the point of cooling; in the case of NGST this would require about 100 W to be rejected at 220 K from the ISIM. Technology demonstrations are planned⁶ for late 1999 to validate operation of a minitarized turbine rotor and self-acting gas bearings with a flight version cooler to be available in 2003.

- 2.) A helium Joule-Thomson cycle cooler⁹ with a mechanical compressor has been demonstrated at 4 K by Rutherford Appleton Laboratory/ Matra Marconi Space Ltd. (RAL/MMS), using a 20 K Stirling upper stage cooler. It is currently being developed as an intermediate cooler¹⁰ (with an 18 K hydrogen-sorption cooler upper stage) for the High-Frequency Instrument on the the ESA Planck mission.
- 3.) Ball Aerospace has a development project with the USAF for a hydrid 10 K cooler that combines a helium J-T stage coupled to a three-stage Stirling cooler¹¹. A new type of rotary vane compressor is being developed for the helium J-T stage. Work has concentrated on developing the He pump, with delivery of a 100 mW at 10 K Engineering Design Model cooler planned by the end of 2000. Taking advantage of the passive thermal cooling of the NGST ISIM to ~35 K, Ball has very recently proposed a two-stage J-T 6 K cooler that would use independent helium and hydrogen J-T loops with separate rotary vane compressors¹². This cooler would reject ~30 W at 220 K, near the ISIM, and 50 W at 270 K on the spacecraft bus.
- 4.) A periodic solid hydrogen sorption cooler¹³ (BETSCE) demonstrated cooling to 10 K during flight operation onboard the Shuttle in May 1996. Pre- and post-flight ground tests produced minimum temperatures down to 9.4 K. Currently, no further development of periodic solid hydrogen sorption cooling is planned.
- 5.) Several organizations have built laboratory liquid hydrogen sorption coolers^{14,15} that operated at cooling temperatures in the range 18 K to 28 K. Currently, JPL is developing a continuous 18-20 K-hydrogen cooler^{16,17} for the Planck mission. An Elegant Breadboard (EBB) cooler is be to operating in the laboratory in 2001 with two flight units to be delivered to Europe by 2004.
- 6.) A Proof-of-Principle (PoP) demonstration of continuous formation of solid hydrogen has been performed¹⁸ in 1993. However, no additional experimental work has been done to demonstrate a continuous solid hydrogen sorption cooler using appropriate hydride compressors.
- 7.) Helium-carbon sorption coolers are well established for periodic operation on the ground¹⁹ and there has been use for short duration rocket flights in space²⁰. The continuous He-charcoal sorption cooler has been demonstrated in the laboratory²¹ but not yet developed for space coolers. The JPL sorption team has developed a basic design²² for 10 mW helium-charcoal coolers to operate in the 5–8 K region as part of this study.

PoP demonstrations are planned for 2000 with demonstration of a complete cooler in 2002 using the 20K EBB hydrogen cooler as the upper stage.

Two-stage sorption coolers offer three attractive options to satisfy the temperature requirements for MIR detectors in the NGST instrument. The low-temperature cooling requirement is presumed to be 10 mW or less. All would use the nominal 18 K liquid hydrogen upper stage cooler being developed for Planck. If the ISIM radiators can reject sufficient heat loads at 30 to 35 K, a substantial reduction in size, mass, and power for the hydride compressor would be realized from the Planck cooler.

1. For cooling to the 8-14 K region, a continuous solid hydrogen metal hydride compressor could be developed. This option requires near term development and demonstration of the integration of the solid hydrogen J-T cryostat and low-pressure (i.e., absorption pressures below 1 torr) sorbent bed compressors.
2. For cooling to the 5–8 K region, the sorption helium-charcoal cooler would provide the least vibrations and should also be scaleable to smaller loads. Although many of the basic principles of this cooler are understood and have been demonstrated, substantial technology development is required to develop a flight-qualified continuous-cycle cooler. This option is described more fully in the next section.
3. For cooling to the 4–8 K region, the mechanical compressor helium J-T cooler from either RAL/MMS or Ball could be integrated to the sorption 20 K cooler. Since the RAL/MMS cooler is to be mounted to the Planck sorption cooler, a combination that will be fully flight qualified by Planck, this is the least expensive, lowest risk, and most expeditious choice. However, potential vibration issues from the mechanical helium compressors remain.

A. 2 Sorption Cooler Design for an NGST MIR Camera

System Level

A cooler system designed to provide base temperatures of 6 to 8 K will consist of two stages – a precooler to reach ~18 K and the final low-temperature stage. Additionally, the precooler will require precooling of the working fluid to less than ~60 K, which will be provided by passive radiation. For the NGST ISIM cooler, the precooling temperature is expected to be 35-40 K. In

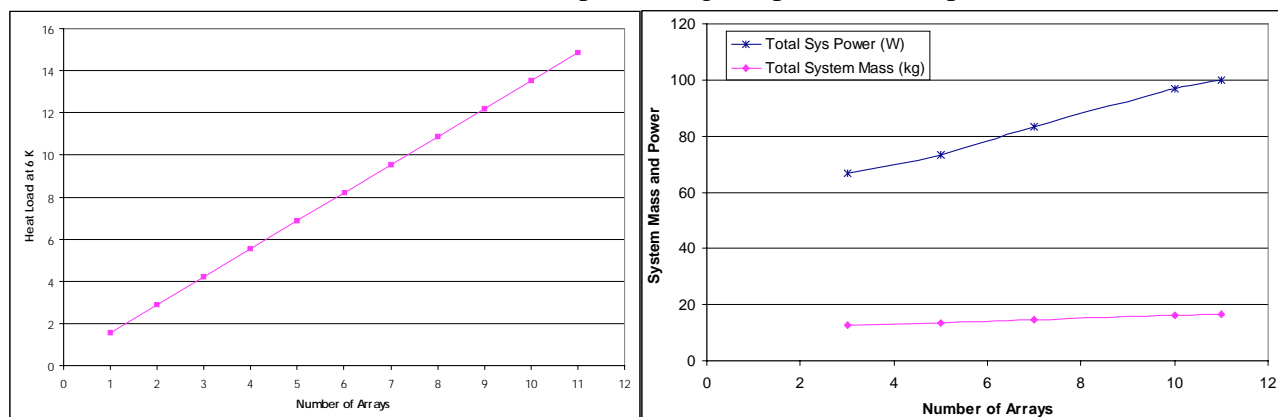


Figure A1. A) The heat lift required at 6 K as a function of the number of detector arrays. B) the total system mass and power as a function of the number of detector arrays.

Table 2. Cooler System Properties for Various 6 K Cooling Loads. For all cases the combined electronics package is assumed to have a mass of 8 kg and an additional power requirement of 10 W. The radiator mass for the hydride cooler system compressors is assumed to be 10 kg.

Heat Lift At 6 K (W)	Charcoal Input Power (W) (at 18 K)	Charcoal Sys Mass (kg)	Hydride Input power (W) (at 270 K)	Hydride System Mass (kg)	Total System Power (W)	Total System Mass (kg)	Passive Cooling requirements (W)	
							At 35 K	At 270 K
0.005	0.43	.56	66.4	12	76.8	30.6	0.44	66.4
0.010	0.81	.95	82.6	13.7	93.4	32.6	0.82	82.6
0.015	1.20	1.34	98.9	15.4	110.1	34.7	1.21	98.9

this section, we present a design for a two-stage continuous sorption cooler system capable of achieving 6 K. Details of each of the stages will be presented separately, along with a summary of the combined system, because either of the stages could be used separately in a combination sorption/mechanical cooler system, as described in above.

The heat load on the low-temperature stage is expected to be < 10 mW, depending on the number and characteristics of the specific detector arrays used in the MIR focal plane. Each array will dissipate approximately 1 mW, with the parasitic heat leak from lead wires (~ 20 /array, plus 16 for temperature monitoring and control) contributing an additional 0.03 mW/array. The heat leak through the supporting structure will also be small, assumed to be 0.16 mW independent of the number of arrays. We expect that seven 1024x1024 arrays would require ~ 10 mW of heat lift at 6 K (including parasitic heat loads to a single support structure from an 18 K precooling stage); the dependence of the cold-stage heat load on the number of detector arrays is shown in Figure A1.

In order to provide 10 mW of heat lift, which would be sufficient for 7 detector arrays, the complete cooler system is expected to weigh 15 kg, requiring a total input power of 83 W, with a cooling efficiency of 8300 watts input power per watt of heat lift at 6 K. Table 2 shows the mass and power requirements of each of the stages, as well as the complete system, for heat lift of 5, 10, and 15 mW at 6 K.

18 K Sorption Cooler

The precooler design favored in this study is a hydrogen-sorption refrigerator, based on the cooler in development for the Planck Surveyor mission^{16,17}. The compressor system is composed of a number of sorbent beds, each of which contain a $\text{LaNi}_{4.8}\text{Sn}_{0.2}$ alloy that is capable of absorbing (and also desorbing) a substantial amount of H_2 . Typical operation temperatures are 270 K for absorption of hydrogen and 470 K

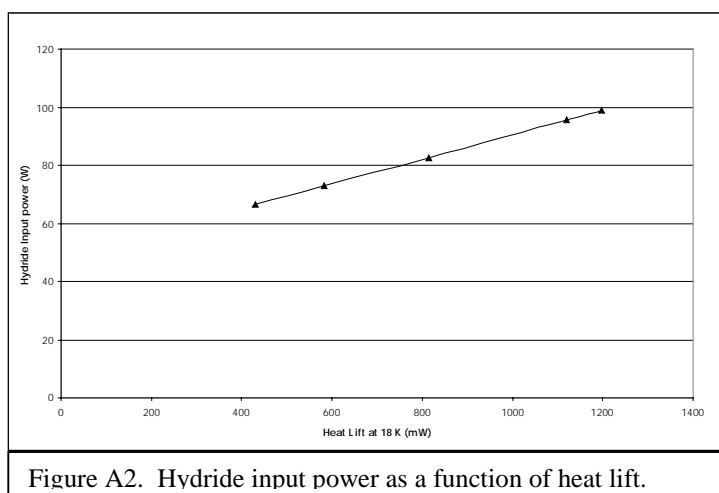


Figure A2. Hydride input power as a function of heat lift.

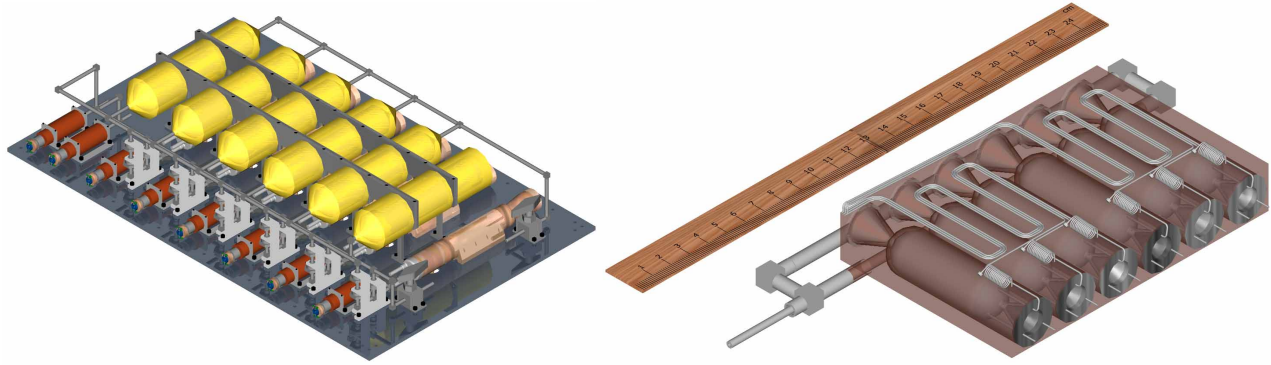


Figure A3. left) Scale drawing of the hydrogen sorption compressor system. The mounting plate is 500 mm x 800 mm x 7 mm. right) Scale drawing of the helium sorption cooler system design. The block containing the compressor elements also contains the Joule-Thomson constriction for the hydride cooler. The hydrogen precooling supply comes through the tube at the lower left. The outer block dimensions are 138 x 92 x 28 mm.

for desorption. The heat-exhausted in cooling from 470 K back to 270 K is dissipated to space by a passive radiator located on the warm side of the spacecraft bus. The compressed hydrogen gas produced by the compressor system is cycled through a Joule-Thomson constriction at the cold head (~ 18 K) to provide precooling for the 6 K cooler and to cool the detector enclosures. The heat removed from the 18 K stage is rejected almost entirely at the 35 K ISIM radiator.

One of the major advantages of the hydrogen-sorption cooler is that it can be easily scaled to meet different requirements by simply changing the length of the compressor beds. The 35-40 K passive precooling provided by NGST will require very small compressor beds – only ~ 10 cm long. This is substantially smaller than the compressors for the Planck system, and a redesign of the mechanical support system will enable further mass and power savings. Input power as a function of heat lift at 18 K is shown in Figure A2. A scale drawing of the Planck Sorption cooler system (after resizing for NGST ISIM requirements) is shown in Figure A3.

For an NGST hydrogen sorption cooler, the hydrogen compressor beds would be located on the warm side of the spacecraft bus and coupled to a passive radiator at ~ 270 K. The temperature oscillations of the 270 K radiator due to cycling of the compressors are expected to be less than ± 1 K at ~ 1 mHz. Gas-gap heat switches completely enclose the compressor beds, shielding the rest of the spacecraft from the 470 K compressors, and are used to couple(decouple) the compressor beds to(from) the radiator during absorption(desorption). The outer shells of the gas gap heat switches are tightly coupled to the radiator and remain at roughly the radiator temperature, even when the compressor beds are at 470 K. The heat switches also function by hydrogen sorption, but with different alloys as the sorbent material. The gas generated by the compressors is transported to (and from) the 18 K cold head through tubes that uncoil during deployment of the ISIM support. The transport tubes also function as a tube-in-tube heat exchanger and are heat-sunk to the radiation shields as they pass through them from the warm side to the cold side. A schematic of the complete system, including the low-temperature stage, is shown in Figure A4. The 18 K cooler can also cool the radiation shielding between the 6 K stage and the 35 K ISIM into its acceptable range.

6 K cooler

The 6 K cooler functions on a principle similar to that of the 18 K cooler, but uses carbon (charcoal) as the sorbent material and helium as the working fluid. Periodic charcoal-helium sorption coolers have been in common laboratory use for many years¹⁹, and have been flown on sounding-rocket flights²⁰. Mechanically driven 10 mW Joule-Thomson coolers have already been developed for flight applications²³, and are baselined as the 4 K cooler for ESA's Planck Surveyor mission. Charcoal-sorption coolers have the advantages of simplicity, easy scalability, and zero vibration because they have no moving parts. Low power, closed cycle operation of a Joule-Thomson cooler driven by charcoal compressors was demonstrated in the laboratory more than 10 years ago by Duband²¹. In such systems helium is physisorbed onto the charcoal, rather than chemisorbed, and the system is not susceptible to degradation of the sorbent materials due to cycling or overheating.

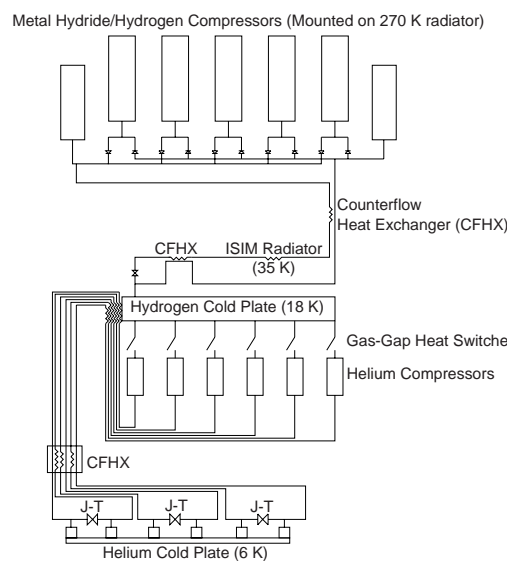


Figure A4. Schematic of the combined cooler system

The basic design of the present cooler model uses charcoal compressor beds that cycle between 80 K for desorption of the helium and 18 K for adsorption. The compressor beds are paired such that each pair of compressors has its own Joule-Thomson constriction that can function for flow in either direction – the gas is driven back and forth between the two compressor beds by alternating which bed is hot and which is cold. Three pairs of beds, operating out of phase, are used to provide continuous cooling, with the three pairs sharing a common counterflow heat exchanger. For camera designs in which there are multiple focal planes that are spatially separated each focal plane can be provided with its own set of compressors. Because the mass and power requirements of the compressors scale nearly linearly with the heat lift, multiple smaller coolers (e.g. two 5 mW coolers) are only slightly more massive and require only slightly more power than a single 10 mW cooler. The main mass penalty is from multiple supporting structures. The main cause of power increase will be due to the parasitic heat loads on the multiple focal planes.

In the present design the helium is always maintained above its critical point ($T_c=5.2$ K, $P_c=0.23$ MPa for ^4He ; $T_c=3.32$ K, $P_c=0.117$ MPa for ^3He), eliminating the need for reservoirs at the cold stage to capture liquid. The design can be readily modified to allow liquid production and capture. Heating and cooling of the beds is done with resistive heaters and gas-gap heat switches, respectively. Sorption driven gas-gap heat switches for the 4 to 10 K temperature range have been flown on sounding rocket flights²⁴. Such heat switches work by alternately filling a narrow space between the compressor bed and the precooling stage with helium to provide thermal conduction between the stage and compressor and removing the gas to provide isolation. Separate small, low power (\sim mW) sorption pumps would be used to fill and evacuate the gas gaps. Figure A4 shows a schematic of the charcoal-helium system coupled to a hydrogen sorption cooler.

Because the system has no moving parts, the failure modes are very limited; the most likely failure is plugging of the Joule-Thomson constriction due to contamination of the helium. This can be avoided by careful processing and handling of the system components during manufacture. By keeping the gas from each pair of beds isolated from the others, the system offers built-in redundancy against failure of a single subsystem. If one pair of beds fails, the timing on the remaining pairs can be adjusted to still provide continuous cooling, at the expense of somewhat larger temperature swings at the cold head. Micromachined check-valves are in development at the University of Twente in the Netherlands²⁵, in cooperation with JPL, that would allow the complete system to operate with only a single Joule-Thomson valve. Check valves would reduce the complexity of the plumbing (making the system topology essentially the same as the hydrogen sorption system) at the possible expense of redundancy. If low-temperature check valves are used, the helium charcoal gas-flow circuit in Figure A4 would be replaced by a circuit much like that of the higher-temperature upper stage.

As shown in Table 2, a 10 mW charcoal helium cooler will require ~815 mW input power, which must be exhausted at the 18 K precooling stage, and would have a total mass of approximately 1 kg. Figure A3 shows preliminary designs of the hydrogen-sorption and helium-sorption compressors, and Figure A5 shows the same design as it might fit into the ISIM. The visible shells of the compressors beds in Figure A3 are the outsides of the gas-gap heat switches, and remain at approximately the temperature of the 18 K cold stage. The charcoal-containing compressor bed volumes are located inside of the gas-gap shells and are thermally isolated so that when a compressor bed is hot (~80 K) and desorbing there is less than 10 mW of heat leak to the 18 K stage.

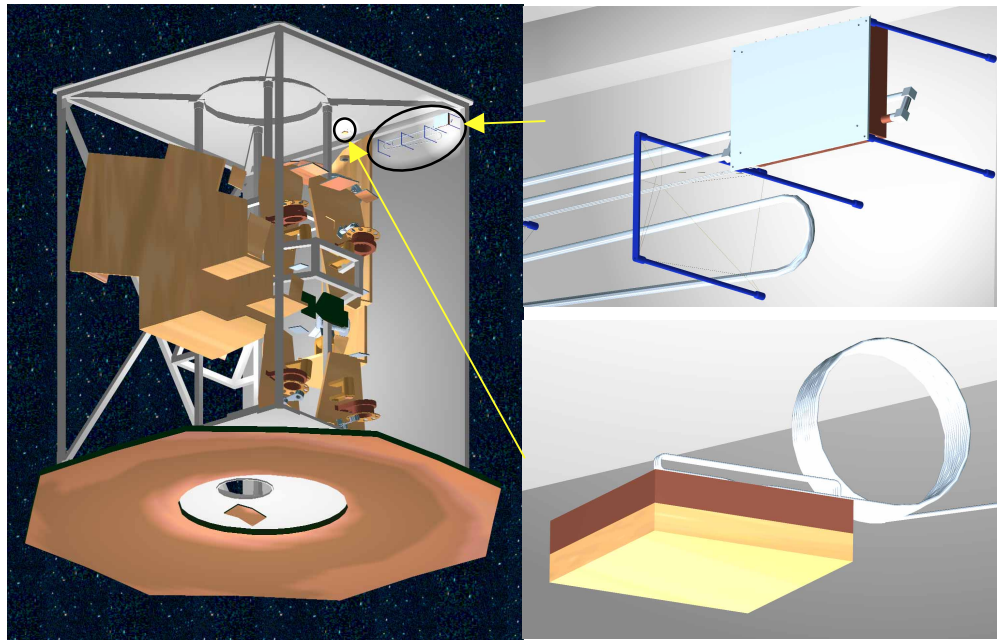


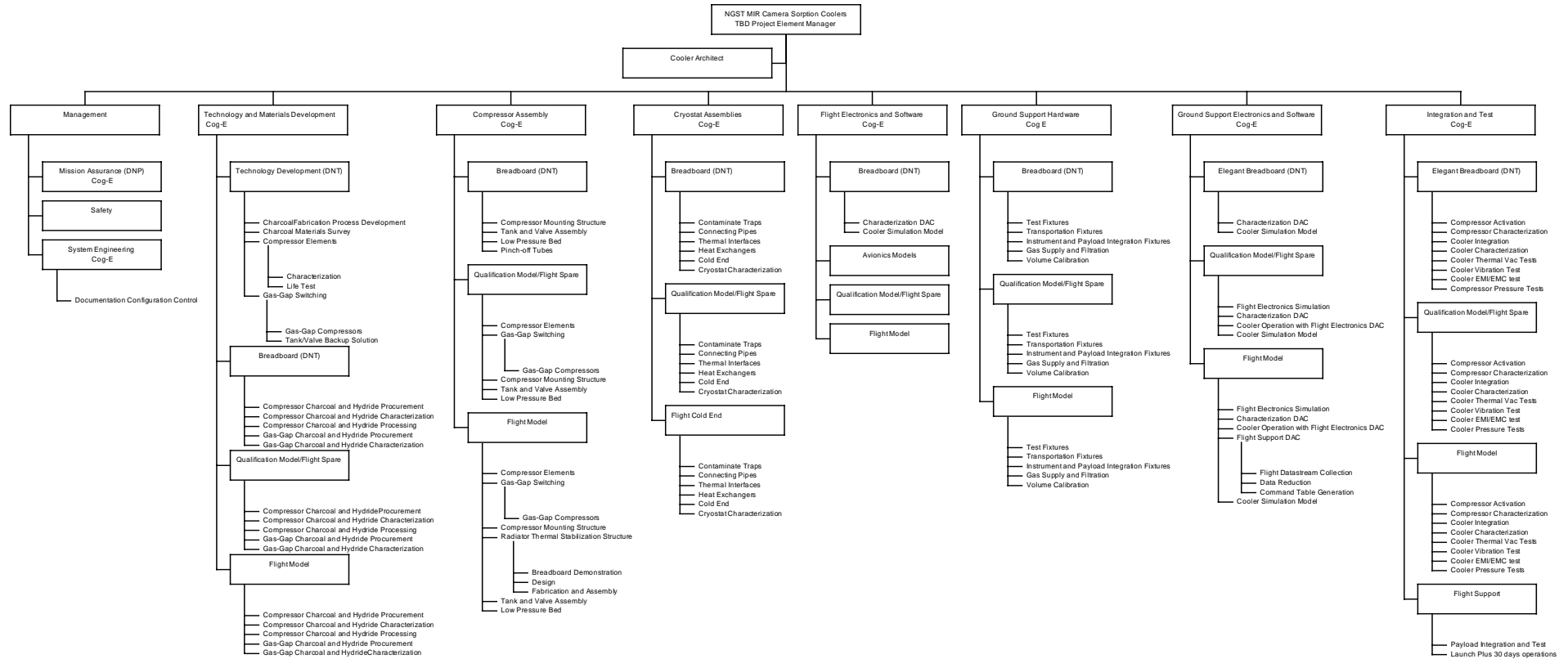
Figure A5. The lefthand figure shows how the charcoal compressor system might mount in the ISIM. The compressor is located in the upper right of the ISIM (inside the ellipse), and the focal plane is located slightly to the left of the compressor (circled). The figures on the right show details at the compressor (upper) and focal plane (lower). The tubes are gas supply and return lines. The yellow area at the focal plane indicates where the detectors would be mounted.

- ¹ D. S. Glaister, M. Donabedian, D. G. T. Curran, and T. Davis, "An Overview of the Performance and Maturity of Long Life Cryocoolers for Space Applications", in *Cryocoolers 10*, edited by R. G. Ross, Jr. (Kluwer Academic/Plenum, New York, 1999), 1-19.
- ² A. H. Orlowska, T. W. Bradshaw, S. Scull, and B. J. Tomlinson, "Progress Towards the Development of a 10 K Closed Cycled Cooler for Space Use", in *Cryocoolers 10*, edited by R. G. Ross, Jr. (Kluwer Academic/Plenum, New York, 1999), 67-76.
- ³ C. Wang, and P. E. Gifford, "A Small-Scale Liquid Helium Plant by Using of 4 K Pulse Tube Cryorefrigerator" presented at Cryogenics Engineering Conference, Montreal, July, 1999; Cryomech Website at URL location: www.cryomech.com/pulsetube.html
- ⁴ G. R. Chandratilleke, et al., "Conceptual Design of Space Qualified 4K Pulse Tube Cryocooler" in *Cryocoolers 10*, edited by R. G. Ross, Jr. (Kluwer Academic/Plenum, New York, 1999), 221-226; N. Yoshimura, et al., "Performance Dependence of 4K Pulse Tube Cryocooler on Working Pressure", in *Cryocoolers 10*, edited by R. G. Ross, Jr. (Kluwer Academic/Plenum, New York, 1999), 227-232.
- ⁵ T. M. Davis, J. Reilly, B. J. Tomlinson, "Air Force Research Laboratory Cryocooler Technology Development", in *Cryocoolers 10*, edited by R. G. Ross, Jr. (Kluwer Academic/Plenum, New York, 1999), 21-32.
- ⁶ W. L. Swift, et al., "Developments in Turbo Brayton Technology for Low Temperature Applications", presented at Space Cryogenics Workshop, Quebec, July, 1999.
- ⁷ J. A. McCormick, W. L. Swift, and H. Sixsmith, "Progress on the Development of Miniature Turbomachines for Low-Capacity Reverse-Brayton Cryocoolers", in *Cryocoolers 9*, edited by R. G. Ross, Jr. (Plenum, New York, 1997), 475-483.
- ⁸ J. Bruning and T. Pilson, "Phillips Laboratory Space Cryocooler Development and Test Program", *Adv. Cryogenic Eng.* **43**, 1651 (1998).
- ⁹ B. G. Jones and D. W. Ramsay, "Qualification of a 4 K Mechanical Cooler for Space Applications" in *Cryocoolers 8*, edited by R. G. Ross, Jr. (Plenum, New York, 1995), 525-535; R. S. Bhatia, A. G. Murray, M. J. Griffin, P. A. R. Ade, T. W. Bradshaw, and A. H. Orlowska, "Integration of a Photoconductive Detector with a 4 K Cryocooler" in *Cryocoolers 9*, edited by R. G. Ross, Jr. (Plenum, New York, 1997), 949-957.
- ¹⁰ A. N. Wombwell, S. R. Scull, B. G. Jones, T. W. Bradshaw, A. H. Orlowska, and C. I. Jewell, "Design and Development of a 20 K/ 4 K Mechanical Cooler", *Proc. 1998 Space Cryogenics Workshop* (In Press); S. R. Scull, B. G. Jones, T. W. Bradshaw, A. H. Orlowska, and C. I. Jewell, "Design and Development of a 4K Mechanical Cooler", in *Cryocoolers 10*, edited by R. G. Ross, Jr. (Kluwer Academic/Plenum, New York, 1999), 513-519.
- ¹¹ R. Levenduski, W. Gully, and J. Lester, "Hybrid 10 K Cryocooler for Space Applications", in *Cryocoolers 10*, edited by R. G. Ross, Jr. (Kluwer Academic/Plenum, New York, 1999), 505-511.
- ¹² R. L. Oonk, D. S. Glaister, and R. Woodruff, Private Communication, August, 1999.
- ¹³ R. C. Bowman, Jr., P. B. Karlmann, and S. Bard, Brilliant Eyes Ten-Kelvin Sorption Cryocooler Experiment (BETSCE) Final Report, JPL Technical Publication 97-14, Issued September 12, 1997.
- ¹⁴ J. A. Jones and P. M. Golben, "Design, Life Testing, and Future Designs of Cryogenic Hydride Refrigeration Systems", *Cryogenics* **25**, 212 (1985).
- ¹⁵ L. Zhang, X. Y. Yu, Y. M. Zhou, G. Ke, F. Zhan, D. Y. Bao, and G. R. Qin, "Development of a 25 K Hydrogen Refrigerator with Hydride Sorption Compressor", *Adv. Cryogenic Eng.* **39**, 1525 (1994).
- ¹⁶ P. Bhandari, R.C. Bowman, R.G.Chave, C.A. Lindensmith, G. Morgante, C. Paine, M. Prina, and L.A.Wade, "Sorption Cryocooler Development for the Planck Surveyor Mission," *Astrophysical Letters and Communications*, (in press)
- ¹⁷ L. A. Wade, et al, "Hydrogen Sorption Cryocoolers for the Planck Mission", *Adv. Cryogenic Engineering* **45** (Submitted).
- ¹⁸ R. C. Longworth and A. Khatri, "Continuous Flow Cryogen Sublimation Cooler", *Adv. Cryogenic Eng.* **41**, 1297 (1996).
- ¹⁹ G. Seidel and P.H. Keesom, *Rev. Scient. Instrum.* **29**, 606 (1958); G.K. White, "Experimental Techniques in Low Temperature Physics," Oxford University Press, (New York) 1993.
- ²⁰ J.J. Bock, L. Duband, M. Kawada, H. Matsuhara, T. Matsumoto, and A.E. Lange, "⁴He refrigerator for space," *Cryogenics* **34**, 635 (1994).
- ²¹ Lionel Duband, Ph.D. Thesis, L'Universite Scientifique, Technologique et Medicale de Grenoble (France 1987)
- ²² C. A. Lindensmith, et al., "Models for Scaleable Helium-Carbon Sorption Cryocoolers", *Adv. Cryogenic Engineering* **45** (Submitted).
- ²³ T.W.Bradshaw and A.H.Orlowska, "The Use of Closed Cycle Coolers on Space Based Observatories," *Space Science Reviews* **74**, 205 (1995)

²⁴M.M. Freund, L. Duband, A.E. Lange, T. Matsumoto, H. Murakami, T. Hirao, and S. Sato, "Design and Flight Performance of a Space Borne ^3He refrigerator for the Infrared Telescope in Space," *Cryogenics* **38**, 435 (1998).

²⁵J. F. Burger, M.C. van der Wekken, E. Berenschot, H.J. Holland, H.J.M ter Brake, H. Rogalla, J.G.E. Gardeniers, and M. Elwenspoek, "High pressure Check Valve for Application in a Miniature Cryogenic Sorption Cooler," preprint (1999); H. Holland, M. ter Brake, H. van Egmond, H. Gardeniers, M. Elwenspoek, H. Rogalla, "Components of a Sorption Based Microcooler", presented at Cryogenics Engineering Conference, Montreal, July, 1999.

NGST Sorption Cryocooler Project Element Work Breakdown Structure



September 2001 Breadboard Charcoal Cooler using GM precooler
 september 2002 Brassboard scaled Hydride cooler
 september 2003 Brassboard cooler

Position	Name	Rate	FY99	selection		CDR		Deliver Flight Inst				Total	Cost(\$K)
				FY'00	FY'01	FY'02	FY'03	FY'04	FY'05	FY '06	FY'07		
				Phase A		Phase C							
PEM/Sys Engr		92	0.0	0.3	0.3	1.0	1.0	1.0	1.0	0.6	0.0	5.2	861
System Architect		95	0.0	0.2	0.5	0.9	0.9	0.9	0.9	0.9	0.0	5.2	878
Tech Devel Eng		75	0.0	1.0	1.0	0.5	0.5	0.0	0.0	0.0	0.0	3.0	441
brassboard Cog E		82	0.0	0.0	0.0	1.1	1.1	0.8	0.5	0.3	0.2	3.9	603
Compressor Cog E		95	0.0	0.0	0.0	1.2	0.6	0.4	0.4	0.2	0.2	3.0	507
Test Cog E		83	0.0	0.0	0.0	0.5	0.5	1.0	1.0	0.5	0.2	3.7	576
Mech Eng/ GSE Hardware Cog		85	0.0	0.0	0.0	1.0	1.0	1.0	0.5	0.1	0.0	3.6	569
Fab Engineer		83	0.0	0.5	0.5	0.5	0.6	0.5	0.2	0.1	0.0	2.9	452
Technician 1		60	0.0	1.0	1.0	2.0	2.0	2.0	2.0	1.0	0.2	11.2	1462
Technician 2		55	0.0	1.0	1.0	2.0	2.0	1.5	0.5	0.0	0.0	8.0	1000
GSE Electronics		75	0.0	0.0	0.0	0.6	0.7	1.2	0.6	0.2	0.0	3.3	485
Flight Electronics		75	0.0	0.0	0.2	1.5	2.0	1.5	1.2	0.5	0.0	6.9	1014
Structural Engr		80	0.0	0.1	0.1	0.3	0.4	0.3	0.2	0.1	0.0	1.5	229
Design Engineer		55	0.0	0.7	0.7	1.3	0.6	0.2	0.0	0.0	0.0	3.5	431
Product Assurance		75	0.0	0.1	0.1	0.3	0.3	0.3	0.3	0.2	0.0	1.6	235
Safety Engr		75	0.0	0.1	0.1	0.2	0.2	0.2	0.2	0.1	0.0	1.1	162
Admin		45	0.0	0.2	0.2	0.5	0.5	0.5	0.5	0.3	0.0	2.7	303
Supervisor Support		90	0.0	0.2	0.2	0.4	0.4	0.4	0.3	0.3	0.0	2.2	360
Total FTE			0.0	5.4	5.9	15.7	14.8	13.7	10.3	5.4	0.8	72.4	10566
Workforce Costs													
DNS/EIS/Phone fees*workforce-K\$/workyear	5		0.0	27.0	29.5	78.5	73.8	68.5	51.5	26.8	4.0	360	
Salary	Rate x Work Years	Salary	0.0	357.6	401.1	1136.9	1067.9	1004.1	761.6	412.6	64.0	5206	
Benefits	Benefit Rate x Salary	0.0424	0.0	15.2	17.0	48.2	45.3	42.6	32.3	17.5	2.7	221	
Hourly Overhead	Hours x Hourly Rate	27.29	0.0	306.5	334.9	891.2	837.3	777.7	584.7	303.7	45.4	4081	
General Burden	Subtotal x Gen Rate	0.052	0.0	35.3	39.2	108.0	101.4	94.9	71.7	38.2	5.8	494	
Total Labor (FY'99\$)			0.0	741.6	822	2263	2126	1988	1502	799	122	10362	
Procurements			0.0	350.0	400	800	900	750	500	200	0	3900	
Travel			0.0	10.0	10	70	70	85	100	50	30	425	
Environmental Test			0.0	0.0	0	50	100	300	300	100	0	850	
Reserves	20%		0.0	220.3	246	637	639	625	480	230	30	3107	
Total Cost (FY'99\$)			0	1322	1478	3819	3835	3747	2882	1378	182	18644	
Inflation factor			0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04		
Caltech 2.2% fee	0.022		0	30	35	93	97	99	79	40	5.08	479	
Real Year Costs			0	1405	1634	4326	4512	4600	3685	1854	235.80	22251	